

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2004, 13:43:06 ; Search time 120 Seconds
(without alignments)
5799.237 Million cell updates/sec

Title: US-09-600-714-41
Perfect score: 1254
Sequence: 1 atgagcttaagtaccgcg.....attggctgttgattttaa 1254

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A-COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B-COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A-COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B-COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS-COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	1185.2	94.5	1384	1	US-08-553-888A-1		Sequence 1, Appli
2	1182	94.3	1466	1	US-08-553-888A-2		Sequence 2, Appli
3	109.6	8.7	1650	4	US-09-461-325-75		Sequence 75, Appl
4	109.6	8.7	1650	4	US-10-012-542-75		Sequence 75, Appl
5	41	3.3	1884	4	US-09-621-976-14663		Sequence 14663, A
6	41	3.3	2094	4	US-09-252-991A-16118		Sequence 16118, A
7	41	3.3	2616	4	US-09-252-991A-16486		Sequence 16486, A
8	41	3.3	6909	4	US-09-199-637A-111		Sequence 111, App
9	40.6	3.2	1422	4	US-09-252-991A-8351		Sequence 8351, Ap
10	40.6	3.2	1503	4	US-09-252-991A-8488		Sequence 8488, Ap
11	40.6	3.2	1761	4	US-09-252-991A-8423		Sequence 8423, Ap
12	40.2	3.2	729	4	US-09-489-039A-5837		Sequence 5837, Ap
13	40	3.2	801	4	US-09-252-991A-6902		Sequence 6902, Ap
14	40	3.2	1410	4	US-09-252-991A-6871		Sequence 6871, Ap
15	40	3.2	1482	4	US-09-252-991A-6850		Sequence 6850, Ap
16	40	3.2	1500	4	US-09-252-991A-6757		Sequence 6757, Ap
17	39.4	3.1	467	2	US-08-476-176B-15		Sequence 15, Appl
18	39.4	3.1	467	3	US-08-127-721A-15		Sequence 15, Appl
19	39.4	3.1	467	3	US-08-485-246A-15		Sequence 15, Appl
20	39.4	3.1	468	2	US-08-476-176B-11		Sequence 11, Appl
21	39.4	3.1	468	3	US-08-127-721A-11		Sequence 11, Appl
22	39.4	3.1	468	3	US-08-485-246A-11		Sequence 11, Appl
23	39	3.1	1564	4	US-09-369-247-25		Sequence 25, Appl
24	38.8	3.1	4403765	3	US-09-103-840A-2		Sequence 2, Appli
25	38.8	3.1	4411529	3	US-09-103-840A-1		Sequence 1, Appli
26	38.2	3.0	2712	4	US-09-919-172-40		Sequence 40, Appl
27	38	3.0	1189	1	US-07-781-034-4		Sequence 4, Appli

28	38	3.0	1189	5	PCT-US92-08328-4	Sequence 4, Appli
29	37.8	3.0	500	4	US-09-370-838-142	Sequence 142, App
30	37.2	3.0	1384	4	US-09-372-422A-17	Sequence 17, Appl
31	36.6	2.9	897	4	US-09-668-680-5	Sequence 5, Appli
32	36.6	2.9	2031	4	US-09-475-515-60	Sequence 60, Appl
33	36.6	2.9	2553	4	US-09-475-515-66	Sequence 66, Appl
34	36.6	2.9	536165	4	US-09-214-808-1	Sequence 1, Appli
35	36.6	2.9	4403765	3	US-09-103-840A-2	Sequence 2, Appli
36	36.6	2.9	4411529	3	US-09-103-840A-1	Sequence 1, Appli
37	36.4	2.9	377	2	US-08-332-766A-1	Sequence 1, Appli
38	36.2	2.9	467	2	US-08-476-176B-17	Sequence 17, Appl
39	36.2	2.9	467	3	US-08-127-721A-17	Sequence 17, Appl
40	36.2	2.9	467	3	US-08-485-246A-17	Sequence 17, Appl
41	36.2	2.9	468	2	US-08-476-176B-13	Sequence 13, Appl
42	36.2	2.9	468	3	US-08-127-721A-13	Sequence 13, Appl
43	36.2	2.9	468	3	US-08-485-246A-13	Sequence 13, Appl
44	36.2	2.9	1941	4	US-09-489-039A-5585	Sequence 5585, Ap
45	36.2	2.9	2004	1	US-08-471-033-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-553-888A-1
; Sequence 1, Application US/08553888A
; Patent No. 5723293
; GENERAL INFORMATION:
; APPLICANT: Huang
; TITLE OF INVENTION: DIAGNOSTIC METHOD AND KIT FOR
; TITLE OF INVENTION: DETERMINING RH BLOOD GROUP GENOTYPE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/553.888A
; FILING DATE: 11/06/95
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Dea, Sean W.
; REGISTRATION NUMBER: 37690
; REFERENCE/DOCKET NUMBER: 454-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1384 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-553-888A-1

Query Match	94.5%	Score 1185.2	DB 1	Length 1384
Best Local Similarity	96.6%	Pred. No. 0		
Matches 1211	Conservative	0	Mismatches 43	Indels 0
Gaps	0			
QY	1	ATGAGCTCTAAGTACCGCGGTCTCTCGGGCGCTCGCTGCGCCCTTGGCCCTTACACTG	60	
DB	31	ATGAGCTCTAAGTACCGCGGTCTCTCGGGCGCTCGCTGCGCCCTTGGCCCTTACACTG	90	
QY	61	GAAGCAGCTCTCATCTCTCTCTCTATTTTATTTTACCACCTATGAGCTTCCCTTAGAGAT	120	

Db	91	GAAGAGCGCTCTCA	TTCTCTCTCTCTCTATTTT	TTTATCCCACTATGACGGCTT	CCTTAGAGGAT	150
Qy	121	CAAAAGGGGCTCGTGGCATCT	ATCAAGTTTGGCCAAAGATCTGA	CCGCGTGAATGGGGCCATT	180	
Db	151	CAAAAGGGGCTCGTGGCATCT	ATCAAGTGGCCAAAGATCTG	ACCCTGATGGCGGCCCTT	210	
Qy	181	GGCTTGGGCTTCTCACTCG	AGTTTCCGGAGACACAGCTGG	AGCAGTGGGCTTCAAC	240	
Db	211	GGCTTGGGCTTCTCACTCA	CTCAAAATTTCCGGAGACACAGCTGG	AGCAGTGGGCTTCAAC	270	
Qy	241	CTCTTCATGCTGGCGCTTGGT	GTGTCAGTGGGCAATCTGCTGG	AGCGCTTCTTGAGCCAG	300	
Db	271	CTCTTCATGCTGGCGCTTGGT	GTGTCAGTGGGCAATCTGCTGG	AGCGCTTCTTGAGCCAG	330	
Qy	301	TTCCCTTCTGGGAAGTGGTCA	NCACACTGTTTCAGTATTCGG	CTGGTGGCCACCATGAGTGT	360	
Db	331	TTCCCTTCTGGGAAGTGGTCA	NCACACTGTTTCAGTATTCGG	CTGGTGGCCACCATGAGTGT	390	
Qy	361	TTGTGCGTGTGATCTCAGTGG	ATGCTGTTGGGGAAGTCAACTTGG	CGCGCAGTTCGGT	420	
Db	391	ATGTGCGTGTGATCTCAGCGG	GTGCTCTTGGGGAAGTCAACTTGG	CGCGCAGTTCGGT	450	
Qy	421	GTGATGTGTGTTGGAGGTG	TCACAGCTTTTAGGCACCTG	AGGATGGTTCATCAGTAATATC	480	
Db	451	GTGATGTGTGTTGGAGGTG	TCACAGCTTTTAGGCACCTG	AGGATGGTTCATCAGTAATATC	510	
Qy	481	TTCAACACAGACTACACATGA	ACATGATGCACATCTACGTGT	TCGCAGCCTATTTTGGG	540	
Db	511	TTCAACACAGACTACACATGA	ACCTCAGGCACCTTCTACGTGT	TCGCAGCCTATTTTGGG	570	
Qy	541	CTGTCTCTGGCGTGGTCC	CAAGCTCTACCCGAGGGAACGG	AGGATAAAGATCAG	600	
Db	571	CTGACTGTGGCTTGGTCTG	CCCAAGCCCTACCCAAGGGAACGG	AGGATAAAGATCAG	630	
Qy	601	ACAGCAACGATACCAAGTGT	CTGCGCATGTGGGCGCCCTCTT	CTTGTGGAATGTTCTGG	660	
Db	631	AGACCAACGATACCAAGTGT	CTGCGCATGTGGGCGCCCTCTT	CTTGTGGAATGTTCTGG	690	
Qy	661	CCAAGTTTCAACTCTGCTCT	GCAGAAAGTCCCAATCGAAAG	GAAGATGCCGTGTTCAAC	720	
Db	691	CCAAAGTCAACTCTGCTCT	GCAGAAAGTCCCAATCGAAAG	GAAGATGCCGTGTTCAAC	750	
Qy	721	ACCTACTATGCTGTAGCAGT	CAGCGTGGTGACGCCATCTC	AGGGTTCATCCTTGGGCTCAC	780	
Db	751	ACCTACTATGCTGTAGCAGT	CAGCGTGGTGACGCCATCTC	AGGGTTCATCCTTGGGCTCAC	810	
Qy	781	CCCCAAGGAAGATCACAAG	ACTTATGTCACAGTGGGTGTGG	CAGAGCGGTGGCT	840	
Db	811	CCCCAAGGAAGATCACAAG	ACTTATGTCACAGTGGGTGTGG	CAGAGCGGTGGCT	870	
Qy	841	GTGGGTACCTCGTGTCACT	GATCCCTTCTCCGTTGGCTTGC	CATGTTGGGCTTGTGTG	900	
Db	871	GTGGGTACCTCGTGTCACT	GATCCCTTCTCCGTTGGCTTGC	CATGTTGGGCTTGTGTG	930	
Qy	901	GCTGGGCTGATCTCCGTG	CGGGGAGCAAAGTACCTGCG	GGGTGTTGTAAACGAGTGTG	960	
Db	931	GCTGGGCTGATCTCCATG	CGGGGAGCAAAGTACCTGCG	GGGTGTTGTAAACGAGTGTG	990	
Qy	961	GGGATTCACCAAGCTCCAT	CATGGGGTACAACCTTCAGT	TGCTGGGCTGCTTGGAGAG	1020	
Db	991	GGGATTCACCAAGCTCCAT	CATGGGGTACAACCTTCAGT	TGCTGGGCTGCTTGGAGAG	1050	
Qy	1021	ATCATCTACATGTGCTGTG	TGCTGTGATACCGTGGAGCGG	GCAATGGCATGATTTGGC	1080	
Db	1051	ATCACCTACATGTGCTGTG	TGCTGTGATACCGTGGAGCGG	GCAATGGCATGATTTGGC	1110	
Qy	1081	TTCCAGTGTCTCTCAGAT	TGGGGAACCTCAGCTTGG	CCCATCGTGATGCTCAAGTCT	1140	
Db	1111	TTCCAGTGTCTCTCAGAT	TGGGGAACCTCAGCTTGG	CCCATCGTGATGCTCAAGTCT	1170	
Qy	1141	GGTCTCTGACAGGTTTG	CTCTAAATCTTAAATATG	GAAGACACCTCATGAGGTTAA	1200	
Db	1171	GGTCTCTGACAGGTTTG	CTCTAAATCTTAAATATG	GAAGACACCTCATGAGGTTAA	1230	

QY	1201	TATTTTGATGACCAAGTTTTCTGGAAGTTTCCCTCAITTTGCTGTGGATTTAA	1254
Db	1231	TATTTTGATGACCAAGTTTTCTGGAAGTTTCCCTCAITTTGCTGTGGATTTAA	1284
RESULT 2			
US-08-553-888A-2			
; Sequence 2, Application US/08553888A			
; Patent No. 5723293			
; GENERAL INFORMATION:			
; APPLICANT: Huang			
; TITLE OF INVENTION: DIAGNOSTIC METHOD AND KIT FOR			
; TITLE OF INVENTION: DETERMINING RH BLOOD GROUP GENOTYPE			
; NUMBER OF SEQUENCES: 3			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Hoffmann & Baron, LLP			
; STREET: 350 Jericho Turnpike			
; CITY: Jericho			
; STATE: New York			
; COUNTRY: USA			
; ZIP: 11753			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage			
; COMPUTER: IBM compatible			
; OPERATING SYSTEM: MS-DOS			
; SOFTWARE: WordPerfect			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/553,888A			
; FILING DATE: 11/06/95			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: O'Dea, Sean W.			
; REGISTRATION NUMBER: 37690			
; REFERENCE/DOCKET NUMBER: 454-5			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (516) 822-3550			
; TELEFAX: (516) 822-3582			
; INFORMATION FOR SEQ ID NO: 2:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1466 nucleotides			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
US-08-553-888A-2			
Query Match 94.3%; Score 1182; DB 1; Length 1466;			
Best Local Similarity 96.4%; Pred. No. 0;			
Matches 1209; Conservative 0; Mismatches 45; Indels 0; Gaps 0;			
QY	1	ATGAGCTCTAAGTACCGCGGTCTGTCGGCGCTGCTGCCCTCTGGGCGCTTAAACATG	60
Db	45	ATGAGCTCTAAGTACCGCGGTCTGTCGGCGCTGCTGCCCTCTGGGCGCTTAAACATG	104
QY	61	GAAGCAGCTCTCATCTCTCCTCTTATTTTTTACCCCACTATGACGCTTCCTTTAGAGGAT	120
Db	105	GAAGCAGCTCTCATCTCTCCTCTTATTTTTTACCCCACTATGACGCTTCCTTTAGAGGAT	164
QY	121	CAAAAGGGGCTCGTGGCATCCTATCAAGTTGGCCAAAGATCGACCGTGATGGCGGCATT	180
Db	165	CAAAAGGGGCTCGTGGCATCCTATCAAGTCGGCCAAAGATCTGACCGTGATGGCGGCCTT	224
QY	181	GGCTTGGGCTTCCTCAGCTCGAGTTTCCGAGACACAGCTGGAGCAGTGTGGCCTTCAAC	240
Db	225	GGCTTGGGCTTCCTCAGCTTCCTACCTCAAAATTCGGAGACACAGCTGGAGCAGTGTGGCCTTCAAC	284
QY	241	CTCTTCATGCTGGCGCTTGCTGTGCGAGTGGGCAATCCTGCTGGACGGCTTCCTTGAGCCAG	300
Db	285	CTCTTCATGCTGGCGCTTGCTGTGCGAGTGGGCAATCCTGCTGGACGGCTTCCTTGAGCCAG	344
QY	301	TTCCCTTCTGGGAGGTGGTCAACACTGTTCACTGATTCGGTGGGCCACCATGAGTGCT	360
Db	345	TTCCCTTCTGGGAGGTGGTCACTACACTGTTTCTGATTCGGTGGGCCACCATGAGTGCT	404

QY	361	TTGTCCGTCGTCTCATCTCAGTGGATCGTCTCTTTGGGAAAGGTCACTCTGGCGCAGTTGGT	420
Db	405	ATGTCCGGTGTCTGATCTCAGCGGTGCTGTCTTTGGGAAAGGTCACTCTGGCGCAGTTGGT	464
QY	421	GTGATGGTGTGGTGGAGGTGACAGCTTTAGGCACCTGAGGATGGTCATCAGTAATAATC	480
Db	465	GTGATGGTGTGGTGGAGGTGACAGCTTTAGGCACCTGAGGATGGTCATCAGTAATAATC	524
QY	481	TTCAACACAGACTACCATGAACATGATGACATCTACGTGTGTTCGACGCTATTTTGGG	540
Db	525	TTCAACACAGACTACCATGAACCTGAGGCATCTCTACGTGTGTTCGACGCTATTTTGGG	584
QY	541	CTGTCTGGGCTGTGTGCTGCGAAAGCCTCTACCCGAGGAGAAACGAGAGTAAGAATCAG	600
Db	585	CTGACTGTGGCTGTGTGCTGCCAAGCCTCTACCCAAGGGAACGAGAGTAAGAATCAG	644
QY	601	ACAGAACGATACCCAGTTTGTCTGCCATGCTGGGCGCCCTCTTCTTGTGGATGTCTTGG	660
Db	645	AGAGCAACGATACCCAGTTTGTCTGCCATGCTGGGCGCCCTCTTCTTGTGGATGTCTTGG	704
QY	661	CCAAGTTTCAACTCTGCTCTCTGCTGAGAAAGTCGATCTGAAAGAGAAATGCCGTGTCAAC	720
Db	705	CCAAGTGTCAACTCTCTCTCTGAGAAAGTCCAATCCAAAGGAAAGAAATGCCATGTCAAC	764
QY	721	ACCTACTATGCTGTAGCAGTCAGCGTGTGTGACAGCCATCTCAGAGTCATCTCTGGCCTCAC	780
Db	765	ACCTACTATGCTGTAGCAGTCAGTGTGTGTGACAGCCATCTCAGAGTCATCTCTGGCCTCAC	824
QY	781	CCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTCGGGTGTTTGGCAGGAGCGGTGGCT	840
Db	825	CCCCAAGGGAAGATCAGCATGACTTATGTGTGCACAGTCGGGTGTTTGGCAGGAGCGGTGGCT	884
QY	841	GTGGGTACTCTGTGTCACTGATCCCTTCTCCGTGGCTTGCATGTGTGCTGGGTCTTGTGTG	900
Db	885	GTGGGTACTCTGTGTCACTGATCCCTTCTCCGTGGCTTGCATGTGTGCTGGGTCTTGTGTG	944
QY	901	GCTGGGCTGATCTCCGTCCGGGGAGCCAAAGTACCTCCGGGGTGTGTGAACCCGAGTGGT	960
Db	945	GCTGGGCTGATCTCCATCCGGGGAGCCAAAGTGCCTTCCGGGTGTGTGTGAACCCGAGTGGT	1004
QY	961	GGGATTCGCCACAGCTCCATCATGGGCTTACAACTTTCAGTTTCTGTGGGTCTGTCTGGAGAG	1020
Db	1005	GGGATTCGCCACATCTCCGTGATGCACCTCCATCTTTCAGTTTCTGTGGGTCTGTCTGGAGAG	1064
QY	1021	ATCATCTACATTTGCTGTCTGGTGTGATACCGTCGGAGCCGGCAATGGCATGATTTGGC	1080
Db	1065	ATCACCTACATTTGCTGTCTGGTGTGATACCGTCGGAGCCGGCAATGGCATGATTTGGC	1124
QY	1081	TTCCAGGTCTCCCTCAGCATTTGGGAACTCAGCTTGGCCATCGTGATAGCTCTCACCTCT	1140
Db	1125	TTCCAGGTCTCCCTCAGCATTTGGGAACTCAGCTTGGCCATCGTGATAGCTCTCACCTCT	1184
QY	1141	GGTCTCCTGCACAGTTTGTCTCTAAATCTTTAAATATGGAAGACACCTCATGAGGCTAAA	1200
Db	1185	GGTCTCCTGCACAGTTTGTCTCTAAATCTTTAAATATGGAAGACACCTCATGAGGCTAAA	1244
QY	1201	TATTTTGAATGACCAAGTTTTCTGGAAGTTTCTCCTCATTTTGGCTGTGTGATTTTAA	1254
Db	1245	TATTTTGAATGACCAAGTTTTCTGGAAGTTTCTCCTCATTTTGGCTGTGTGATTTTAA	1298

RESULT 3

RESOLUTION 3
US-09-461-325-75

US 05-401-323-73
; Sequence 75, Application US/09461325A

; Patent No. 6475753

; FACILE NO. 6473733
; GENERAL INFORMATION:

APPLICANT: Ruben et al.

; AFFILICANI: RUBEN ET AL.;
; TITLE OF INVENTION: 94 Human Secreted Proteins

FILE OF INVENTION: 34
FILE REFERENCE: PZ029P1

FILE REFERENCE: FZ029FI
CURRENT APPLICATION NUMBER: US/09/461.325A

; CURRENT AFFILIATION NUMBER: US/1
; CURRENT FILING DATE: 1999-12-14

; CONCURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418

770 CCTTGGCTACCCCCCAAGGAAGATCAGCAAGACTTATGTGCACAGTGCGGTGTTGGCAG 829

Db 1826 CTTGGTACGCGCCGCTGACCTGGACCCCTGAGTGTGACGACGCGGAGATTTG 1767
QY 830 GAGGCTGTGTGGTACTGTGTGACATGATCCCTTCTCGTGGCTTGCCATGGTGC 889
Db 1766 GCGTCACGCCCGGCTGCGACCTTCTCAAGATTTCCCGCGATCCCGGTACCAAG 1707
QY 890 TGGGCTTTGTGGCTGGCTGATCTCGTGGGGAGCCAAAGTACCTGCGGGGTGTGTA 949
Db 1706 TGATTTCCGAGCCTGGCGACCAACATCTGTGTGGCGGCTGGCTGCGCTGACCGGAC 1647
QY 950 ACCGAGTGTGGGATTTCCCAAGCTCCATCATGGGCTACAACTTCAGCTTGTGGGTC 1009
Db 1646 GCAGCAACCCGCCAAGCCCAAGCGTGCCTGCCGGGCGCGCTGGCAGCGGTGGCT 1587
QY 1010 TCGTTGGAGATATCATATGTGCTGCTGTGTGATACCGTTCGAGCGGCAATG 1069
Db 1586 CGCTGCGCGGTTCTATCTGCTGTTGTTGATGTGGCGCAGACCTCGGTGCGCACCTACT 1527
QY 1070 GCATGATTGGCTTCCAGGTCTCTCCTCAGC 1098
Db 1526 ACATGAAGGATCTGCTGCTTACCAGGC 1498

RESULT 7

US-09-252-991A-16486
; Sequence 16486, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16486
; LENGTH: 2616
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16486

Query Match 3.3%; Score 41; DB 4; Length 2616;
Best Local Similarity 45.3%; Pred. No. 0.051;
Matches 149; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

QY 770 CTTTGGCTACCCCAAGGAGATCAGCAAGATTTATGTGACAGTGGGTGTGGCAG 829
Db 203 CTTGGTACGCGCCGCTGCGAGCTGGGACCCCTGAGCTGGACGACGCGGAGATTTG 262
QY 830 GAGGCTGTGGTACTCTGCTGCTACCTGATCCCTTCTCGGTGGCTTGCCATGGTGC 889
Db 263 GGTGAGCCCGGCTGCGACCTTCTCAAGATTTCCCGCGATCCGCCGTACCAAG 322
QY 890 TGGGCTTGTGGCTGGGTGATCTCGTGGGGAGCCAAAGTACCTGCGGGGTGTGTA 949
Db 323 TGATTTCCGAGCCTGGCGCACCAACATCTGTGTGGCGGTGGCGTGGCTGACCGGAC 382
QY 950 ACCGAGTGTGGGATTTCCCAAGCTCCATCATGGGCTAACAATTCAGCTTGTGGTGC 1009
Db 383 GCAGCAACCCGCCAAGCCCAAGCTGCGCTCCCGGGCGCGCTGGCAGCGGTGGCT 442
QY 1010 TCGTTGGAGATCATCTACATTTGCTGCTGTGCTTATACCGTTCGAGCGGCAATG 1069
Db 443 CGCTGCGCGGTTCTATCTGCTGTTGTTGATGCTGGCGCAGACCTCGGTGCGCACCTACT 502
QY 1070 GCATGATTGGCTTCCAGGTCTCTCCTCAGC 1098
Db 503 ACATGAAGGATCTGCTGCTTACCAGGC 1498

RESULT 8

US-09-199-637A-111
; Sequence 111, Application US/09199637A
; Patent No. 635411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: RastSEQ for Windows Version 4.0
; SEQ ID NO 111
; LENGTH: 6909
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-111

Query Match 3.3%; Score 41; DB 4; Length 6909;
Best Local Similarity 45.3%; Pred. No. 0.091;
Matches 149; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

QY 770 CTTTGGCTACCCCAAGGAGATCAGCAAGATTTATGTGACAGTGGGTGTGGCAG 829
Db 1497 CTTGGTACGCGCCGCTGCGAGCTGGGACCCCTGAGCTGACGACGCGGAGATTTG 1556
QY 830 GAGGCTGTGGTACTCTGCTGCTACCTGATCCCTTCTCGGTGGCTTGCCATGGTGC 889
Db 1557 GGTGCGACGCCAGGCTGCGACCTTCTCAAGATTTCCCGCGATCCGCCGTACCAAG 1616
QY 890 TGGGCTTGTGGCTGGGTGATCTCGTGGGGAGCCAAAGTACCTGCGGGGTGTGTA 949
Db 1617 TGATTTCCGAGCCTGGCGCACCAACATCTGTGTGGCGGTGGCTGACCGGAC 1676

QY 950 ACCGAGTGTGGGATTTCCCAAGCTCCATCATGGGCTACAACTTCAGCTTGTGGTGC 1009
Db 1677 GCAGCAACCCGCCAAGCCCAAGCGTGGCTGCCCGGGCGCGCTGGCAGCGGTGCGCT 1736
QY 1010 TCGTTGGAGATCATCTACATTTGTGTGCTGTGCTGTGATACCGTTCGAGCGGCAATG 1069
Db 1737 CGCTGCGCGGTTCTATCTGCTGTTGTTGATGCTGGCGCAGACCTCGGTGCGCACCTACT 1796
QY 1070 GCATGATTGGCTTCCAGGTCTCTCCTCAGC 1098
Db 1797 ACATGAAGGATCTGCTGCTTACCAGGC 1825

RESULT 9

US-09-252-991A-8351/c
; Sequence 8351, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

Search completed: August 25, 2004, 17:30:45
Job time : 130 secs

Score	Match	Length	DB	ID	Description
1252.4	99.9	1354	17	US-10-672-764A-35	Sequence
365.8	29.2	487	10	US-09-918-995-30694	Sequence
175.4	14.0	1792	13	US-10-302-172-362	Sequence
165.4	13.2	1351	16	US-10-074-978A-41	Sequence
165.4	13.2	1440	9	US-09-949-145-5	Sequence
165.4	13.2	1952	9	US-09-949-145-1	Sequence
165.4	13.2	1952	13	US-10-342-887-1635	Sequence
165.4	13.2	1952	13	US-10-172-118-1635	Sequence
150.2	12.0	1497	9	US-09-949-145-6	Sequence
150.2	12.0	2097	9	US-09-949-145-2	Sequence
148.8	11.9	2415	16	US-10-104-047-730	Sequence
144.6	11.5	505	13	US-10-027-632-282039	Sequence
144.6	11.5	505	13	US-10-027-632-282040	Sequence
144.6	11.5	505	16	US-10-027-633-282039	Sequence

QY	241	CTCTTCATGCTGCGCTGCTGGTGGAGTGGGCAATCCTGCTGACGGCTTCCTGAGCCAG	300
Db	241	CTCTTCATGCTGCGCTGCTGGTGGAGTGGGCAATCCTGCTGAGCGCTTCCTGAGCCAG	300
QY	301	TTCCCTCTCGGAGTGGTGCATCACACTGTTTCAGTATTCGGCTGCCACCATGAGTGCT	360
Db	301	TTCCCTCTCGGAGTGGTGCATCACACTGTTTCAGTATTCGGCTGCCACCATGAGTGCT	360
QY	361	TTGTCGGTGTGATCTCAGTGGATGCTGTTCTGGGGAAGTCAACTTGGCGCAGTTGGTG	420
Db	361	TTGTCGGTGTGATCTCAGTGGATGCTGTTCTGGGGAAGTCAACTTGGCGCAGTTGGTG	420
QY	421	GTGATGGTGTGCTGAGGTGACAGCTTTTAGGCAACTGAGGATGTCATCAGTAATATC	480
Db	421	GTGATGGTGTGCTGAGGTGACAGCTTTTAGGCAACTGAGGATGTCATCAGTAATATC	480
QY	481	TTCAACACAGCTACCAATGAATGATGATGATGATGATGATGATGATGATGATGATGATG	540
Db	481	TTCAACACAGCTACCAATGAATGATGATGATGATGATGATGATGATGATGATGATGATG	540
QY	541	CTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	600
Db	541	CTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	600
QY	601	ACAGCAACGATACCCAGTTTGTCTGCCATGCTGGGCGCCCTCTTCTTGTGGATGTTCTGG	660
Db	601	ACAGCAACGATACCCAGTTTGTCTGCCATGCTGGGCGCCCTCTTCTTGTGGATGTTCTGG	660
QY	661	CCAGTTTCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	720
Db	661	CCAGTTTCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	720
QY	721	ACCTACTATGCTGTAGCAGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	780
Db	721	ACCTACTATGCTGTAGCAGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	780
QY	781	CCCCAGGAGATACAGAGACTTATGTCACAGTGGGCTGCTGCTGCTGCTGCTGCTGCTG	840
Db	781	CCCCAGGAGATACAGAGACTTATGTCACAGTGGGCTGCTGCTGCTGCTGCTGCTGCTG	840
QY	841	GTGGGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	900
Db	841	GTGGGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	900
QY	901	GCTGGGTGATCTCCGTGGGGAGCCAAAGTACCTGCTGGGGTGTGTAACCGAGTGGTG	960
Db	901	GCTGGGTGATCTCCGTGGGGAGCCAAAGTACCTGCTGGGGTGTGTAACCGAGTGGTG	960
QY	961	GGGATTCCTCCACAGCTCCATCATGGCTACAACTTCAGCTTGTGGGTCTGCTTGGAGAG	1020
Db	961	GGGATTCCTCCACAGCTCCATCATGGCTACAACTTCAGCTTGTGGGTCTGCTTGGAGAG	1020
QY	1021	ATCATCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1080
Db	1021	ATCATCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1080
QY	1081	TTCCAGGTCTCTCAGCAATTTGGGAACTCAGTCTGGCCATCTGCTGCTGCTGCTGCTG	1140
Db	1081	TTCCAGGTCTCTCAGCAATTTGGGAACTCAGTCTGGCCATCTGCTGCTGCTGCTGCTG	1140
QY	1141	GGTCTCCTGACAGTGTGCTCCTAAATCTTAAATATGGAAGACACCTCATGAGGCTAAA	1200
Db	1141	GGTCTCCTGACAGTGTGCTCCTAAATCTTAAATATGGAAGACACCTCATGAGGCTAAA	1200
QY	1201	TATTTTGTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGAATTTTAA	1254
Db	1201	TATTTTGTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGAATTTTAA	1254
RESULT 2			
US-09-918-995-30694			
; Sequence 30694, Application US/09918995			

QY	706	AATGCCGTGTTCAACACCTACTATGCTGTAGCAGTCAAGCTGCTGACAGCCATCTCAGGG	765
Db	56	ATTCCCGTGTTCACACCTACTATGCTGTAGCAGTCAAGCTGCTGACAGCCATCTCAGGG	115
QY	766	TCATCTCTTGGCTCACCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTGGGTGTTG	825
Db	116	TCATCTCTTGGCTCACCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTGGGTGTTG	175
QY	826	GCAGGAGCGTGGCTGTGGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	885
Db	176	GCAGGAGCGTGGCTGTGGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	235
QY	886	GTGCTGGTCTTGTGGCTGGGCTGATCTCGTGGGGGAGCCAAAGTACCTGCCGGGTGT	945
Db	236	GTGCTGGTCTTGTGGCTGGGCTGATCTCGTGGGGGAGCCAAAGTACCTGCCGGGTGT	295
QY	946	TGTAACCGAGTGTGGGATTTCCCCACAGCTCCATCATGGGCTACAACTTCAGCTTGGTG	1005
Db	296	TGTAACCGAGTGTGGGATTTCCCCACAGCTCCATCATGGGCTACAACTTCAGCTTGGTG	355
QY	1006	GGTCTGCTGGAGAGATCATCTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1065
Db	356	GGTCTGCTGGAGAGATCATCTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	415
QY	1066	AATGCCATG 1074	
Db	416	AATGCCATG 424	
RESULT 3			
US-10-302-172-362			
; Sequence 362, Application US/10302172			
; Publication No. US20040053250A1			
; GENERAL INFORMATION:			
; APPLICANT: Tang, Y. Tom			
; APPLICANT: Xue, Aidong J.			
; APPLICANT: Drmanac, Radoje T.			
; TITLE OF INVENTION: No. US20040053250A1el Arginine-rich Protein-like Nucleic Acids			
; TITLE OF INVENTION: Polypeptides			
; FILE REFERENCE: 803.1CNCIP			
; CURRENT APPLICATION NUMBER: US/10/302,172			
; CURRENT FILING DATE: 2002-11-21			
; PRIOR APPLICATION NUMBER: US 10/225,251			
; PRIOR FILING DATE: 2002-08-20			
; PRIOR APPLICATION NUMBER: PCT US02/05095			
; PRIOR FILING DATE: 2002-03-05			
; PRIOR APPLICATION NUMBER: US 09/799,451			

Publication No. US20030073623A1

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

FILE REFERENCE: 20411-756

CURRENT APPLICATION NUMBER: US/09/918,995

CURRENT FILING DATE: 2001-07-30

PRIOR APPLICATION NUMBER: US/09/235,076

PRIOR FILING DATE: 1999-01-20

NUMBER OF SEQ ID NOS: 38054

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 30694

LENGTH: 487

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

LOCATION: (1)...(487)

OTHER INFORMATION: n = A,T,C or G

US-09-918-995-30694

Query Match 29.2%; Score 365.8; DB 10; Length 487;

Best Local Similarity 99.5%; Pred. No. 5.3e-107;

Matches 367; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

; PRIOR FILING DATE: 2001-03-05
 ; NUMBER OF SEQ ID NOS: 950
 ; SOFTWARE: pt_Fl_genes Version 2.0
 ; SEQ ID NO 362
 ; LENGTH: 1792
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (32)..(1354)
 US-10-302-172-362

Query Match 14.0%; Score 175.4; DB 13; Length 1792;
 Best Local Similarity 49.4%; Pred. No. 3.2e-45;
 Matches 455; Conservative 0; Mismatches 466; Indels 0; Gaps 0;

QY	141	CTATCAAGTTGGCCAAAGATCTGACCGTGTATGCGGCCATTGGCTTGGCGTTCCTCACTC	200
DB	211	CTTACCAAGCTTCCAGAGCGTGCATGCCATGTCTTTCGTGGGCTTTGACTTCCCTCATGGT	270
QY	201	GAGTTTCOGAGACACAGCTGGAGCAGTGTGGCCCTTCAACCTCTTTCATGCTGGCGCTTGG	260
DB	271	CTTCTCTGCAGCGTTACGGCTTCAGCAGCGTGGCTTCACTTCTCTCTGGCGCTTGC	330
QY	261	TGTGCAAGTGGCAATCCTCTGAGAGGCTTCTGAGCCAGTTCCTCTGGAAGGTGGT	320
DB	331	CTTGCAGTGTGCACACTGTCCAGGGCTTCTTCCACTCTTCCACGGTGGCCACATCCA	390
QY	321	CATCACACTGTTTCAGTATTCGGCTGGCCACCATGAGTCTTGTGCGTGTGATCTCAGT	380
DB	391	TGTTGGCGTGGAGAGCATGATCAATGCTGACTTTTGTGGGGGCGGTGCTCATCTCTT	450
QY	381	GGATGCTGCTTTGGGGAAGGTCAACTTGGCGCAGTGTGGTGTGATGTGTGTGGAGGT	440
DB	451	TGTTGGCGTGGAGAGCATGATCAATGCTGACTTTTGTGGGGGCGGTGCTCATCTCTT	510
QY	441	GACAGCTTTTAGGCAACCTGAGATGGTATCAGTAAATATCTTCAACACAGACTACCAAT	500
DB	511	GGTGTGTTTGGCATCAATGAGTTTGTGTCTCTTCACTCTCTGGGGGTGAGAGATGCCGG	570
QY	501	GAACATGATGCATCTACGTCTTCGACGCTATTTTGGGCTGTCTGTGGCTGTGTCCT	560
DB	571	AGGCTCATGACTATCCACACTTTTGGTGCCTACTTGGGCTGTGCTTTCGGGGTCT	630
QY	561	GCCAAAGCTCTACCGAGGGAACGAGAGATAAGATCAGACAGCAACGATACCAGTTT	620
DB	631	GTACAGGCCCAAGCTGGAGAGCAAGCAACCGCCAGGGCTCCGTCTACCATTCAGACCT	690
QY	621	GTCTGCATGCTGGGGCCCTCTTCTTGTGATGTCTTGGCAAGTTTCAACTCTGTCT	680
DB	691	CTTGGCCATGATTTGGGACCATCTTCTGTGATCTTCTGGCTTAGCTTCAATGTGCACT	750
QY	681	GCTGAGAAGTCCAAATCGAAAGAAAGATGCGCTGTTCAACACCTACTATGCTGTAGCAGT	740
DB	751	CACAGCGCTGGGGCTGGGCAGCATCGGACGCCCTCAACATACTACTCTCTGGCTGC	810
QY	741	CAGCGTGTGACAGCCATCTCAGGCTCATCTTGGCTCACCCCAAGGGAAGATCAGCAA	800
DB	811	CAGCACCTTGGCACTTTGCTTGTCTGTCAGCCCTTGTAGGGGAAGATGGAGGCTTGACAT	870
QY	801	GACTTATGTCACAGTGGGGTGTGGCAGGAGGCTGTGCTGGGTACCTGCTGTCACCT	860
DB	871	GGTCCACATCCAAAATGACAGCGCTGGCTGGAGGGGTGTGGTGGGGACCTCAAGTGAAT	930
QY	861	GATCCCTTCTCCGTGGCTTGCCATGCTGCTGGTCTTGTGGCTGGGCTGATCTCCGCTCG	920
DB	931	GATGTCGACACCTTGGGGCTCTGGAGCTGTGGCTTCTTGTGCTGGGACTGTCTCCAGCT	990
QY	921	GGGAGCCAAAGTACCTGCCGGGTGTGTAAACGAGTGTCTGGGGATTCCTCCACAGCTCCAT	980
DB	991	GGGTACAGTGTCTTTCACGCCCATCTTGTGAATCAAAATTCAAAGTCCACAGACATGTGG	1050
QY	981	CATGGGCTCAACTTCAGCTTGTGGGTCTGCTGTTGGAGAGATCAATCTACATGTGTGCT	1040

Db 1051 AGTCCACAACCTCCATGGATGCCGGGGTCTCTGGGGCCCTCTCTGGGGGTCCTTGTGGC 1111
 QY 1041 GGTGCTTGATACCGTCGGAGC 1061
 Db 1111 TGGACTTGGCCACCCATGAAGC 1131

RESULT 4
 US-10-074-978A-41/c
 ; Sequence 41, Application US/10074978A
 ; Publication NO. US20040010119A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Leite, Mario
 ; APPLICANT: Spytek, Kimberly A
 ; APPLICANT: Guo, Xiaojia (Sasha)
 ; APPLICANT: Fernandes, Elma
 ; APPLICANT: Li, Li
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Liu, Xiahong
 ; APPLICANT: Casman, Stacie
 ; APPLICANT: Boldog, Ferenc
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Blalock, Angela
 ; APPLICANT: Ballinger, Robert
 ; APPLICANT: Vernet, Corine
 ; APPLICANT: Tchernev, Velizar T
 ; APPLICANT: Malyankar, Uriel M
 ; APPLICANT: Gusev, Vladimir
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Mezes, Peter S
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Heyes, Melvin P
 ; APPLICANT: Herrman, John
 ; APPLICANT: Pena, Carol E A
 ; APPLICANT: Shimkets, Richard A
 ; APPLICANT: Taupier Jr, Raymond J
 ; APPLICANT: Moore, No. US20040010119A11le
 ; APPLICANT: Shenoy, Suresh
 ; APPLICANT: Edinger, Shlomit
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Stone, Dave
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Peyman, John
 ; APPLICANT: Smithson, Glennad
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 ; FILE REFERENCE: 21402-269
 ; CURRENT APPLICATION NUMBER: US/10/074,978A
 ; CURRENT FILING DATE: 2003-01-07
 ; PRIOR APPLICATION NUMBER: 60/268,221
 ; PRIOR FILING DATE: 2001-02-12
 ; PRIOR APPLICATION NUMBER: 60/335,109
 ; PRIOR FILING DATE: 2001-10-31
 ; PRIOR APPLICATION NUMBER: 60/312,284
 ; PRIOR FILING DATE: 2001-08-14
 ; PRIOR APPLICATION NUMBER: 60/268,496
 ; PRIOR FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/276,703
 ; PRIOR FILING DATE: 2001-03-16
 ; PRIOR APPLICATION NUMBER: 60/330,293
 ; PRIOR FILING DATE: 2001-10-18
 ; PRIOR APPLICATION NUMBER: 60/322,127
 ; PRIOR FILING DATE: 2001-11-21
 ; PRIOR APPLICATION NUMBER: 60/280,899
 ; PRIOR FILING DATE: 2001-04-02
 ; PRIOR APPLICATION NUMBER: 60/310,797
 ; PRIOR FILING DATE: 2001-08-08
 ; PRIOR APPLICATION NUMBER: 60/268,646
 ; PRIOR FILING DATE: 2001-02-14
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 547
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 41


```
; LENGTH: 1351
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-074-978A-41

Query Match      13.2%; Score 165.4; DB 16; Length 1351;
Best Local Similarity 49.1%; Pred. No. 4.6e-42; Indels 0; Gaps 0;
Matches 439; Conservative 0; Mismatches 456;

QY 141 CTATCAAGTTGGCCAAAGATCTGACCGTGATGGCGCCATTGGCTTGGGCTTCTCCACCTC 200
Db 1175 CTACCCAAAGCTTCCAGGACGTGCACGTGATGTCCTTGGTGGCTTCCGCTTCTCATGAC 1116

QY 201 GAGTTTCCGGAGACACAGCTGGAGAGTGGCGCTTCAACCTTTCATGCTGGCGCTTGG 260
Db 1115 TTTCCTGCGACGTGACGGCTTCAAGCCGCGTGGGCTTCAACTTCTGTGGCAGCTTCGG 1056

QY 261 TGTGCAAGTGGCAATCTCTGCTGACAGGCTTCTTGAGCCAGTTCCTTCTGGGAAGTGGT 320
Db 1055 CATCCAGTGGCGCTGCTCATGACAGGCTGGTTCCTTCAAGACCGCTACATCGT 996

QY 321 CATCACACTCTTCAGTATTCGGCTGGCCACCATGAGTGTCTTGTGGTGTGATCTCAGT 380
Db 995 CGTGGCGGTGGAGAACCTCATCAACGCTGACTTCTGGTGGCTCTCTGTCTGGCGTTCCT 936

QY 381 GATGCTGTCTTGGGAGAGTCAACTTGGCGCAGTTGGTGGTGTGATGGTGGTGGAGT 440
Db 935 TGGGGCAGTTCTGGGTAAAGTCAGCCCATTCAGTCTGCTCATGATGACTTCTTCCAGT 876

QY 441 GACAGCTTTTAGGCAACCTGAGGATGTCATCAGTAATATCTTCAACACAGACTACCACAT 500
Db 875 GACCTCTTCGCTGTGAATGAGTTTCACTTCTTAACTTAACTTAAAGTGAAGATGCAGG 816

QY 501 GAAATGATGACATCTAGTGTTCGAGCCATATTTTGGGCTGTCCTGAGTGGTGGCTT 560
Db 815 AGGCTCATGACCATCACACATTTGGCGCTACTTTTGGGCTCACAGTGAACCGGATCCT 756

QY 561 GCCAAGCTCTACCCAGGGAACGGAGGATTAAGATCAGACAGCAACGATACCCAGTTT 620
Db 755 CTACCGACGCAACCTAGAGAGAGCAAGAGGAGAGAGAGAAATCTGTGACAGTCCGACCT 696

QY 621 GTCTGCCATGTGGCGGCTCTTCTTGTGATGTTCTGGCCAAAGTTTCAACTCTGCTCT 680
Db 695 CTTTGGCATGATTGGCACCTCTTCTCTGTGATGTAATGACCCAGCTTCAACTCAGCCAT 636

QY 681 GCTGAGAAGTCCAAATCGAAAGGAAGATGCGGTGTTCAACACCTACTATGCTGTAGAGT 740
Db 635 ATCCTACCATGGGACACAGCCAGCACCCGAGCGCCATCAACACCTTACTGCTCTTGGCAGC 576

QY 741 CAGGTGTCACAGCCATCTCAGGTCATCTTGGCTCACCCCAAGGGAAGATCAGCAA 800
Db 575 CTGGTGCTTACCTCGGTGGCAATATCAAGTGGCTCTGACNAGAGGCGAGCTGGACAT 516

QY 801 GACTTATGTGCACAGTGGGTGTTGGAGGAGCGGTGGCTGTGGGTACCTGCTGTCACCT 860
Db 515 GGTGCACATCCAGAAATCCACGCTCGCAGGAGGGGTGGCGTGGGTACCGTGTGAGAT 456

QY 861 GATCCCTTCTCGGTGGCTTGCATGTCGTGGGTCTTGTGGCTGGGTGATCTCCGTGG 920
Db 455 GATGCTCATGCTTACGCTGGCTCATCATCGGCTTCTGTGGCGCATCATCTCCACCT 396

QY 921 GGGAGCCAAAGTACCTGCGGGGTGTTGTAACCGAGTGTGGGATTCGCCACAGTCCAT 980
Db 395 GGGTTTGTATACCTGACCCATTCCTGAGTCCGGCTGCACATCCAGGACACATGTGG 336

QY 981 CATGGCTTACAACTTCACTGCTGGGTCTGCTTGGAGAGATCATCTACATTTG 1035
Db 335 CATTAACAATCTGCATGGCAATTCCTGGCATCATAGGCGGCATCGTGGGTGCTGTG 281
```

RESULT 5
US-09-949-145-5
; Sequence 5, Application US/09949145

```
; Patent No. US20020055622A1
; GENERAL INFORMATION:
; APPLICANT: New York Blood Center
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyco
; FILE REFERENCE: Docket 454-31
; CURRENT APPLICATION NUMBER: US/09/949,145
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/230660
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 1440
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-949-145-5

Query Match      13.2%; Score 165.4; DB 9; Length 1440;
Best Local Similarity 49.1%; Pred. No. 4.8e-42; Indels 0; Gaps 0;
Matches 439; Conservative 0; Mismatches 456;

QY 141 CTATCAAGTTGGCCAAAGATCTGACCGTGATGGCGCCATTGGCTTGGGCTTCTCCACCTC 200
Db 177 CTACCCAAAGCTTCCAGGACGTGCACGTGATGTCCTTGGGCTTCAACTTCTGTGGCAGCTTCGG 236

QY 201 GAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCAACCTTCTCATGTGGCGTGG 260
Db 237 TTTCCTGACGCTACGGCTTCAAGCCGCTGGGCTTCAACTTCTGTGGCAGCTTCGG 296

QY 261 TGTGCAAGTGGCAATCTCTGCTGGACGGCTTCTTGAGCAGTTCCTTCTGGGAAGTGGT 320
Db 297 CATCCAGTGGCGCTGCTCATGACGGCTTGTTCCTTCAAGACCGCTACATCGT 356

QY 321 CATCACACTCTTCAGTATTCGGCTGGCCACCATGAGTGTCTTCTGCGTGTGATCTCAGT 380
Db 357 CGTGGCGCTGGAGAACTCATCAACGCTGACTTCTGGTGGCTCTGTCTGGCTGGCTT 416

QY 381 GATGCTGTCTTGGGGAAGGTCAACTTGGCGCAGTGTGGTGTGATGATGCTGTGGTGGAGT 440
Db 417 TGGGGCAGTCTTGGTAAAGTCAGCCCATTCAGCTGCTCATCATGACTTCTTCTCCAAGT 476

QY 441 GACAGCTTTAGGCAACCTGAGGATGTCATCAGTATATCTTCAACACAGACTACCACAT 500
Db 477 GACCTCTTTCGCTGTGAATGAGTTCATCTTCTTAACTCTAAAGTGAAGATGACAGG 536

QY 501 GAACATGATGCACATCTACGTTTCGACGCTATTTTGGCTGTCTGCGCTGCTGCT 560
Db 537 AGGCTCATGACCATCCACATTTGGCGCTACTTTGGGCTCACAGTGCACCGGATCCT 596

QY 561 GCCAAAGCTCTACCCGAGGAAACGGAGGATAAAGATCAGACAGCAACGATACCCAGTTT 620
Db 597 CTACCGACGCAACCTAGAGCAGAGCAAGGAGAGACAGAAATCTGTGTACAGTGGGACCT 656

QY 621 GTCTGCCATGTGGGCGCTCTTCTTGTGATGTTCTGGCCAAAGTTTCAACTCTGCTCT 680
Db 657 CTTTGGCATGATTGGCACCTCTTCTGTGATGTACTGTGGCCAGCTTCAACTCAGCCAT 716

QY 681 GCTGAGAAGTCCAAATCGAAAGGAGAAATGCGGTTCCTTCAACACCTACTATGCTGTAGCAGT 740
Db 717 ATCTACCATGGGAGACAGCAGCCAGCCAGCGCCATCAACCTACTGCTCTTGGCAGC 776

QY 741 CAGCGTGGTGACAGCCATCTCAGGTCATCTTGGCTCACCCCAAGGGAAGATCAGCAA 800
Db 777 CTGGTCTTACCTCGGTGGCAATATCCAGTGGCTTCAAGAGGAGGAGCTGGACAT 836

QY 801 GACTTATGTGCACAGTGGGTGTGGAGGAGGCTGCTGGTGGTACCTGCTGTCACCT 860
Db 837 GGTGCACATCCAGAAATCCAGCTCGCAGGAGGGGTGGCCGCTGGGTACCGTGTGAGAT 896

QY 861 GATCCCTTCTCGGTGGCTTGCATGTCGTGGGTCTTGTGGCTGGGTGATCTCCGTCGG 920
Db 897 GATGCTCATGCTTACGCTGGCTTCAATCGGCTTCTGTGCGGCATCATCTCCACCT 956
```



```

; APPLICANT: New York Blood Center
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyco
; FILE REFERENCE: Pocket 454-31
; CURRENT APPLICATION NUMBER: US/09/949,145
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/230660
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-949-145-6

Query Match      12.0%; Score 150.2; DB 9; Length 1497;
Best Local Similarity 47.5%; Pred. No. 4e-37;
Matches 446; Conservative 0; Mismatches 493; Indels 0; Gaps 0;

QY  97  CACTATGAGCTTCCTTAGAGGATCAAAAGGGGCTCGTGGCATCTTCCACCTAGGATTTGCATACCTA 156
Db  136  CGCAAGAAATCTCCAGGACGTTGAGACAGTTTACTATGCGTACCGGAGCTTCCAG 195
QY  157  GATCTGACCTGATGCGGCGCAFTGGCTTGGCTTCTCCTCACTCGAGTTTCCGGAGACAC 216
Db  196  GATGTACAGCCATGCTCTTGGTGGCTTCGGCTTCTCTATGACCTCTCTGAGCGCTAC 255
QY  217  AGCTGAGCAGTGTGGCTTCAACCTCTTCATGCTGGCGCTTGGTGTGAGTGGGCAATC 276
Db  256  GGCCTTCAGCGCTGTAGGCTTCACTTCTGCTGGCAGCTTTCGGCATCCAGTGGGCACTG 315
QY  277  CTGCTGGAGCGCTTCTCTGAGCAGTTCCTCTCTGGAAGGTGGTGTATCACAATGTCAGT 336
Db  316  CTCATGACGGATGTTCCATTACTTTGAGAAAGGCCACATTTGCTGAGCGTCGAGAAC 375
QY  337  ATTGCGCTG3CCACCAATGATGTTTGTGCGTGTCTGATCTCAGTGGATGCTGTCTTGGG 396
Db  376  ATCATCAAGCTGACTTCTGTGTGGCATCTTCTGTGTGGCTTCGGGCGAGTTCTAGGC 435
QY  397  AAGGTCACCTTGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 456
Db  436  AAGGTCAGCCCGATGCGCTGCTCATTAATGACCTTCTTCCAACTGACTCTCTTCACAGT 495
QY  457  CTGAGGATGCTCATCAGTAATATCTTCAACAGACTACCAATGATGATGATGATGATGATG 516
Db  496  ATGAGTTCATCTCTGAACTGTATAGGCAAGAGATGAGGGGCTCTATGACCAATC 555
QY  517  TACGTTTGGAGCCTATTGGGCTGTCTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 576
Db  556  CACACATTTGGGCGCTACTTTGGGCTCAAGTGTGATCTCTTACCGAATAACCTG 615
QY  577  GAGGAACGGAGATAAAGATCAGACAGCAAGATACCCAGTTTGTGTGCCATGTGTCAGATC 636
Db  616  GATCAGAGCAAGCAGACAGAGCTCAGTGTATACCACTCGGACCTTTTGGCCATGATGGC 675
QY  637  GGCCTCTTCTGTGATGTTCTGGCCAGTGTTCAACTCTGCTCTGCTGAGAGTCCAATC 696
Db  676  ACCCTCTTCTGTGATATATGCGCCAGTTTCAATTCAGCCAGTTCCTTCCACGGAGAT 735
QY  697  GAAAGGAAGATGCGGTGTTCAACACCTACTATGCTGTAGCAGTCAAGCGTGTGTGACGCG 756
Db  736  GCCAGCACCGAGCAGCGCTCAATACCTACCTCTCTCTTGGCAGCGAGTGTCTAACCA 795
QY  757  ATCTCAGGGTCACTCTTGCTCACCCCAAGGAAGATCAGCAAGACTTATGTGACAGT 816
Db  796  GTGACAGTATCAGTAATGTATACAAAGAAAGGCAAGTTGGATGATGGTGCATCCAGAT 855
QY  817  GCGGTGTTGGCAGGAGCGGTGCTGTGGTACTCTGTGCTACCTGATCCCTTCTCCGTGG 876
Db  856  GCCAGCTTGCAGGTGGGGTGGGTGGGCAAGCTGCGGAGATGATGCTCACACCTTAC 915
QY  877  CTTGCCATGCTGGGTCTTGTGGCTGTGGCTGTATCTCCGTGCGGGGAGCCCAAGTACCTG 936

```

```

Db  916  GGGGCTCTCATCGTGGGTTCTTCTCGGCAATTTTCTCCACCTAGGATTTGCATACCTA 975
QY  937  CCGGGGTGTTAAACCGAGTGTCTGGGATTTCCCCACAGCTCCCATCATGGCTCAAACTTC 996
Db  976  ACGCCATTCCTGGAGTCCCGCTTCGCATCCAGGACACATGGGCAATTCACAACTGCAC 1035
QY  997  AGCTTCTGGTCTGCTTGGAGAGATCATCTACATTTGTG 1035
Db  1036  GGCATTCCTGGCATCATAGCGGCATTTGTGGTGTGTG 1074

RESULT 10
US-09-949-145-2
; Sequence 2, Application US/09949145
; Patent No. US20020055622A1
; GENERAL INFORMATION:
; APPLICANT: New York Blood Center
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyco
; FILE REFERENCE: Pocket 454-31
; CURRENT APPLICATION NUMBER: US/09/949,145
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/230660
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2097
; TYPE: DNA
; ORGANISM: Mus musculus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AF193810
; DATABASE ENTRY DATE: 1999-12-22
; RELEVANT RESIDUES: (1)..(2097)
US-09-949-145-2

Query Match      12.0%; Score 150.2; DB 9; Length 2097;
Best Local Similarity 47.5%; Pred. No. 4.8e-37;
Matches 446; Conservative 0; Mismatches 493; Indels 0; Gaps 0;

QY  97  CACTATGAGCTTCCTTAGAGGATCAAAAGGGGCTCGTGGCATCTTCCATCAAGTTGGCCAA 156
Db  256  CGCAAGAAATCTCCAGGACGTTGAGAACAGATTTTACTATCGCTACCCGAGCTTCAG 317
QY  157  GATCTGACCGTGTATGCGGCGCAATGGCTTGGGCTTCTCCTCACCTCGAGTTTCGGAGACAC 216
Db  318  GATGTACACGGCAATGGTCTTCTGTGGCTTCGGCTTCTCTCATGACCTTCTCTCAGCGCTAC 377
QY  217  AGCTGAGCAGTGTGGCTTCAACCTCTTCTATGCTGGCGCTTGGTGTGAGTGGGCAATC 276
Db  378  GGCCTTCAGCGCTGTAGGCTTCAACTTCTGCTGGCAGCTTTGGCATCCAGTGGGCACTG 437
QY  277  CTGCTGAGCGGCTTCTGAGCCAGTTCCTTCTGGAAGGTGGTCTATCACATGTTCACT 336
Db  438  CTCATGCGAGGATGTTTCCATTACTTTGAAGAGCCACATTTGCTGAGCGTCAGAAC 497
QY  337  ATTGCGTGGGCAACATGAGTGTCTTGTGCGTGTGATCTCAGTGGATGCTGTCTTGGGG 396
Db  498  ATCATCCAAGCTGACTTCTGTGTGGCATCTTCTGTGTGGCTTCTGCGGCGAGTTCCTAGGC 557
QY  397  AAGGTCACCTTGGCGGAGTGGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 456
Db  558  AAGGTCAGCCCGATGCGAGCTGCTCATATGACCTTCTTCCAAAGTACTCTCTTCCAGTGT 617
QY  457  CTGAGGATGTCATCAGTAATATCTTCAACACAGACTACCAATGAAATCATCATGACATC 516
Db  618  AATGAGTTCATCTCTCCATCTGATAGGCAAGAGATGAGGGGCTCTATGACCAATC 677
QY  517  TACGTGTGGCAGCTATTTTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 576
Db  678  CACACATTTGGCGCTACTTTGGGCTTCACAGTGTGAGCTTGGATCTCTACCGAAAAACCTG 737
QY  577  GAGGGAACGGAGGATAAGATCAGACAGCAAGATACCCAGTTTGTGTGTGTGTGTGTGTGT 636

```

Db 738 GATCAGCAAGACAGACAGACAGCTCAGTGTACCACTCGAACCTTTTGGCCATGATTGGC 797
 QY 637 GCCCTCTCTTGTGGATGTTCTGGCAAGTTTCAACTCTGCTCTGCTGAGAGTCAATC 696
 Db 798 ACCCTCTCTTGTGGATATATCTGCCCAAGTTTCAATTCAGCCAGTTTCTTCCACGGAGAT 857
 QY 697 GAAAGGAAGATGCCGTGTTCAACACCTACTATGCTGTAGCATGTCAGCGTGGTGACAGCC 756
 Db 858 GCCCAGACCGAGACGCCCTCAATACCTACTCTCTCTTGGCAGCGAGTGTGTAAACACA 917
 QY 757 ATCTCAGGGTCACTCTTGGCTACCCCAAGGAAAGATCAGCAAGACTTATGTGCACAGT 816
 Db 918 GTGACAGTATCCAGTATGTPACACAAGAGGCAAGTTGGATATGTTGCACATCCAGAAT 977
 QY 817 GCGGTGTGACAGAGCGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 876
 Db 978 GCCACGTTGACAGTGGGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1037
 QY 877 CTGTCATGTTGCTGGGTCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 936
 Db 1038 GCGCTCTCATCTGGGTCTTCTGCGGCAATTTCTCCACCTAGGATTTGCATACCTA 1097
 QY 937 CCGGGTGTGTTAAACCGAGTGTGGGATTTCCCAAGCTTCCATCATGGGCTCAACTTC 996
 Db 1098 ACGCATTCCTGAGTCCCGCTTGCATCCAGGACACATGTGGCAATTCACAACCTGCAC 1157
 QY 997 AGCTTGTGGTCTGCTTGGAGAGATCATCTACATGTG 1035
 Db 1158 GGCATTCCTGGCATCATAGCGCGCAATTTGGGTGCTGTG 1196

RESULT 11
 US-10-104-047-730/c
 ; Sequence 730, Application US/10104047
 ; Publication No. US20030236392A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: No. US20030236392A1 full length cDNA
 ; FILE REFERENCE: HI-A0105
 ; CURRENT APPLICATION NUMBER: US/10/104,047
 ; PRIOR FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER:
 ; PRIOR FILING DATE:
 ; NUMBER OF SEQ ID NOS: 4096
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 730
 ; LENGTH: 2415
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-104-047-730

Query Match 11.9%; Score 148.8; DB 16; Length 2415;
 Best Local Similarity 65.9%; Pred. No. 1.5e-36;
 Matches 216; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
 QY 758 TCTCAGGTCATCTCTGGCTCACCCTCAAGGAAAGATCAGCAAGACTTATGTGCACAGT 817
 Db 1875 TCTCTCTACCTTGTCTTCTTACCCACAGCTATTTCTTTCGACACTTATGTGCACAGT 1816
 QY 818 CGGTGTTGGCAGGAGCGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 877
 Db 1815 CGGTGTTGGCAGGAGCGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1756
 QY 878 TTGCCATGCTGGTCTTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 937
 Db 1755 TTGCCATGCTGGTCTTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1696
 QY 938 CGGGTGTGTAACCGAGTGTGGGATTTCCCAAGCTTCCATCATGGGCTTACAACTTCA 997
 Db 1695 CGGTAAAGAACTAGACAACCTTCTCTGTTGGCTGAAGGCGCAGCAGCAGCTGGG 1636
 QY 998 GCTTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1057

Db 1635 ACCTGATGGCCCACTGTGCAGTGCACAGCTAGGAGGTTGTGGCGCATTTCTCTTA 1576
 QY 1058 GAGCGGCAATGTCATGATTGCTTCCA 1085
 Db 1575 TTGGCTTCAACGCCTAGTGGGATCCA 1548
 RESULT 12
 US-10-027-632-282039
 ; Sequence 282039, Application US/10027632
 ; Publication No. US20020198371A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 282039
 ; LENGTH: 505
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-282039

Query Match 11.5%; Score 144.6; DB 13; Length 505;
 Best Local Similarity 86.9%; Pred. No. 1.4e-35;
 Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
 QY 758 TCTCAGGTCATCTCTGGCTCACCCTCAAGGAAAGATCAGCAAGACTTATGTGCACAGT 817
 Db 241 TCTCTCTACCTTGTCTTCTTACCCACAGCTATTTCTTTCGACACTTATGTGCACAGT 300
 QY 818 CGGTGTTGGCAGGAGCGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 877
 Db 301 CGGTGTTGGCAGGAGCGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 360
 QY 878 TTGCCATGTCGGTCTTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 937
 Db 361 TTGCCATGTCGGTCTTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 420
 QY 938 CGG 940
 Db 421 CGG 423

RESULT 13
 US-10-027-632-282040
 ; Sequence 282040, Application US/10027632
 ; Publication No. US20020198371A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006

```
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 282040
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-282040

Query Match      11.5%; Score 144.6; DB 13; Length 505;
Best Local Similarity 86.9%; Pred. No. 1.4e-35;
Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 758 TCTCAGGGTCATCCTTGGCTCACCCTCCAGGGAAGATCAGCAAGACTTATGTGCACAGTG 817
Db 241 TCTCTCTACCTTGCCTTCTTACCCACACGCTATTTCTTTGGCAGACTTATGTGCACAGTG 300

QY 818 CGGTGTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 877
Db 301 CGGTGTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 360

QY 878 TTGCCATGGTCTGGTCTTGTGGCTGGGTGATCTCCGTGCGGGGAGCAAGTACTCTGC 937
Db 361 TTGCCATGGTCTGGTCTTGTGGCTGGGTGATCTCCGTGCGGGGAGCAAGTACTCTGC 420

QY 938 CGG 940
Db 421 CGG 423

RESULT 14
US-10-027-632-282039
; Sequence 282039, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 282039
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-282040

Query Match      11.5%; Score 144.6; DB 13; Length 505;
Best Local Similarity 86.9%; Pred. No. 1.4e-35;
Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 758 TCTCAGGGTCATCCTTGGCTCACCCTCCAGGGAAGATCAGCAAGACTTATGTGCACAGTG 817
Db 241 TCTCTCTACCTTGCCTTCTTACCCACACGCTATTTCTTTGGCAGACTTATGTGCACAGTG 300

QY 818 CGGTGTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 877
Db 301 CGGTGTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 360

QY 878 TTGCCATGGTCTGGTCTTGTGGCTGGGTGATCTCCGTGCGGGGAGCAAGTACTCTGC 937
Db 361 TTGCCATGGTCTGGTCTTGTGGCTGGGTGATCTCCGTGCGGGGAGCAAGTACTCTGC 420

QY 938 CGG 940
Db 421 CGG 423
```

```
US-10-027-632-282039

Query Match      11.5%; Score 144.6; DB 16; Length 505;
Best Local Similarity 86.9%; Pred. No. 1.4e-35;
Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 758 TCTCAGGGTCATCCTTGGCTCACCCTCCAGGGAAGATCAGCAAGACTTATGTGCACAGTG 817
Db 241 TCTCTCTACCTTGCCTTCTTACCCACACGCTATTTCTTTGGCAGACTTATGTGCACAGTG 300

QY 818 CGGTGTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 877
Db 301 CGGTGTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 360

QY 878 TTGCCATGGTCTGGTCTTGTGGCTGGGTGATCTCCGTGCGGGGAGCAAGTACTCTGC 937
Db 361 TTGCCATGGTCTGGTCTTGTGGCTGGGTGATCTCCGTGCGGGGAGCAAGTACTCTGC 420

QY 938 CGG 940
Db 421 CGG 423

RESULT 15
US-10-027-632-282040
; Sequence 282040, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 282040
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-282040

Query Match      11.5%; Score 144.6; DB 16; Length 505;
Best Local Similarity 86.9%; Pred. No. 1.4e-35;
Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 758 TCTCAGGGTCATCCTTGGCTCACCCTCCAGGGAAGATCAGCAAGACTTATGTGCACAGTG 817
Db 241 TCTCTCTACCTTGCCTTCTTACCCACACGCTATTTCTTTGGCAGACTTATGTGCACAGTG 300

QY 818 CGGTGTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 877
Db 301 CGGTGTGGCAGGAGCGGTGGCTGTGGTACCTCGTGTACCTGATCCCTTCTCCGTGGC 360

QY 878 TTGCCATGGTCTGGTCTTGTGGCTGGGTGATCTCCGTGCGGGGAGCAAGTACTCTGC 937
Db 361 TTGCCATGGTCTGGTCTTGTGGCTGGGTGATCTCCGTGCGGGGAGCAAGTACTCTGC 420

QY 938 CGG 940
```

Db |||
 421 CGG 423

Search completed: August 25, 2004, 18:55:15
Job time : 625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2004, 05:21:23 ; Search time 522 seconds

(without alignments)

10205.441 Million cell updates/sec

Title: US-09-600-714-41

Perfect score: 1254

Sequence: 1 atgagctctaagtaaccgcg.....atttgctgttgattttaa 1254

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002s:*
- 7: Geneseqn2003as:*
- 8: Geneseqn2003bs:*
- 9: Geneseqn2003cs:*
- 10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1254	100.0	1254	2	AAX86522
2	1252.4	99.9	2811	9	ADE09785
3	365.8	29.2	487	8	ACH43482
4	177	14.1	1805	6	ACH43482
5	175.4	14.0	1792	6	ABX12333
6	165.4	13.2	1351	7	ACD20392
7	165.4	13.2	1853	3	AAC64286
8	165.4	13.2	1949	5	AH25763
9	165.4	13.2	1952	6	ABK49215
10	165.4	13.2	1968	7	ABX76257
11	165.4	13.2	1970	3	AAC76344
12	159.4	12.7	1679	6	ABK12334
13	152.4	12.2	3706	4	AH18728
14	150.2	12.0	2098	6	ABK49216
15	148.8	11.9	2415	9	ADB62576
16	147.8	11.8	399	9	ADB62576
17	139.8	11.1	486	4	AAI11615
18	139.8	11.1	486	4	ABA53307
19	139.8	11.1	486	4	AAI32911
20	139.8	11.1	486	4	ABA42886
21	139.8	11.1	486	4	ABA23081
22	139.8	11.1	486	4	AAK27012
23	139.8	11.1	486	4	AAK01568

C	24	139.8	11.1	486	4	ABS26596
C	25	139.8	11.1	486	5	AAI01542
C	26	139.8	11.1	486	6	ABS01595
C	27	135	10.8	1188	9	ADE07444
C	28	120	9.6	1571	6	ABL90445
C	29	113.8	9.1	1840	4	ABL02133
C	30	111.8	8.9	123	4	AAI20829
C	31	111.8	8.9	123	4	ABA65900
C	32	111.8	8.9	123	4	AAI46064
C	33	111.8	8.9	123	4	ABA48015
C	34	111.8	8.9	123	4	ABA32986
C	35	111.8	8.9	123	4	AAK40045
C	36	111.8	8.9	123	4	AAK14315
C	37	111.8	8.9	123	4	ABS39627
C	38	111.8	8.9	123	5	AAI06540
C	39	111.8	8.9	123	6	ABS14127
C	40	109.6	8.7	1650	3	AAZ97083
C	41	109.6	8.7	1650	7	ADA56347
C	42	109.6	8.7	1650	7	ADA40187
C	43	109.6	8.7	1650	8	ACH66712
C	44	109.6	8.7	1650	9	ADC73726
C	45	109.6	8.7	1650	9	ADD37696

ALIGNMENTS

RESULT 1

AAX86522
ID AAX86522 standard; cDNA; 1254 BP.

AC AAX86522;

DT 04-OCT-1999 (first entry)

DE cDNA sequence of the prevalent allele of the Rhd gene.

KW Allele; Rhesus D antigen; Rhd; weak D phenotype; blood transfusion; ss.

OS Homo sapiens.

EH Key Location/Qualifiers
FT CDS 1..1254
FT /*tag= a

PN WO9937763-A2.

PD 29-JUL-1999.

PF 18-DEC-1998; 98WO-EP008319.

PR 23-JAN-1998; 98EP-00101203.

PA (DRKB-) DRK BLUTSPENDEDIENST BADEN WUERTEMBERG.

PI Flegel WA, Wagner FF;

DR WPI; 1999-469127/39.

DR P-PSDB; AAY24056.

PT Nucleic acid sequences correlated with Rhesus weak D phenotype, useful for screening blood from donors and recipients for transfusion methods.

PS Disclosure; Fig 2; 64pp; English.

CC The present sequence represents the prevalent allele of the Rhesus D (Rhd) antigen gene. The specification describes a Rhd contributing to or indicative of the weak D phenotype, where the Rhd polynucleotide carries at least one missense mutation as compared to the wild-type Rhd, in its transmembrane and/or intracellular regions, especially in amino acid positions 2-16, 114-149, 179-225 or/and 267-397, with the proviso that the D antigen does not carry a single missense mutation leading to a F223V or T283I substitution. The probes and antibodies are useful in the

methods for detection of weak D phenotypes. Red blood cells, from probands, are useful for the assessment of the affinity, avidity and/or reactivity of monoclonal anti-D antibodies, polyclonal anti-D antisera or of anti-globulin or anti-human-globulin antisera. Detecting the presence of the RhD associated with weak D phenotype is useful for determining that a patient in need of a blood transfusion is to be transfused with RhD negative blood from a donor. Alternatively, testing for weak D phenotype RhD in the blood of a donor is useful for determining whether the donor blood should be excluded for transfusion to patients having CC wild type RhD or weak D types, other than that of the donor weak D type XX

Sequence 1254 BP; 241 A; 322 C; 351 G; 340 T; 0 U; 0 Other;

Query Match 100.0%; Score 1254; DB 2; Length 1254;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAGCTCTAAGTACCGCGGTCTGTCGCGCTGCTGCCCTCTGGGCGCTTAAACACTG 60
1 ATGAGCTCTAAGTACCGCGGTCTGTCGCGGTCTGTCGCGCTGCTGCCCTCTGGGCGCTTAAACACTG 60
61 GAAGCAGCTCTCAATCTCTCTCTCTATTTTACCCACTATGACGCTTCTTTAGAGGAT 120
61 GAAGCAGCTCTCAATCTCTCTCTCTATTTTACCCACTATGACGCTTCTTTAGAGGAT 120
121 CAAAGGGGCTCGTGGATCTATCAAGTTGGCCAGATCTGACCGTGAAGCGGCCATT 180
121 CAAAGGGGCTCGTGGATCTATCAAGTTGGCCAGATCTGACCGTGAAGCGGCCATT 180
181 GGCTTGGCTTCTCACTCGAGTTTCCGGAGACACAGCTGGAGCAGTGGCCCTTCAAC 240
181 GGCTTGGCTTCTCACTCGAGTTTCCGGAGACACAGCTGGAGCAGTGGCCCTTCAAC 240
241 CTCTTCATGCTGGCGCTTGGTGTGACGTGGGCAATCTGCTGGACGCTTCTTGAGCCAG 300
241 CTCTTCATGCTGGCGCTTGGTGTGACGTGGGCAATCTGCTGGACGCTTCTTGAGCCAG 300
301 TTCCCTTCTGGGAAGTGGTCAACACACTGTTCAGTATTCGGCTGGCCACCATGAGNGCT 360
301 TTCCCTTCTGGGAAGTGGTCAACACACTGTTCAGTATTCGGCTGGCCACCATGAGNGCT 360
361 TTGTCGCTGTGATCTCAGTGGATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGGTG 420
361 TTGTCGCTGTGATCTCAGTGGATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGGTG 420
421 GTGATGCTGTGCTGGAGGTGACAGCTTTAGGCAACTGAGGATGGTCATCAGTAATATC 480
421 GTGATGCTGTGCTGGAGGTGACAGCTTTAGGCAACTGAGGATGGTCATCAGTAATATC 480
481 TTCAACACAGACTACCAACATGAAATGATGACATCTACGTTTCGGAGCCCTATTTTGGG 540
481 TTCAACACAGACTACCAACATGAAATGATGACATCTACGTTTCGGAGCCCTATTTTGGG 540
541 CTGTCCTGGCGCTGGTCTGCCAAAGCTCTACCCGAGGGAACGGAGGATAAAGATCAG 600
541 CTGTCCTGGCGCTGGTCTGCCAAAGCTCTACCCGAGGGAACGGAGGATAAAGATCAG 600
601 ACAGCAAGCATACCCAGTTGCTGCGCATGCTGGCGGCCCTCTTCTTGTGATGTTCTGG 660
601 ACAGCAAGCATACCCAGTTGCTGCGCATGCTGGCGGCCCTCTTCTTGTGATGTTCTGG 660
661 CCAAGTTTCACTCTGCTCTGAGAGTCCCAATCGAAAGGAAGTCCGCTGTTCAAC 720
661 CCAAGTTTCACTCTGCTCTGAGAGTCCCAATCGAAAGGAAGTCCGCTGTTCAAC 720
721 ACCTACTATGCTGTAGCAGTCAAGCTGTGACAGCATCTCAGGGTCACTCTTGGCTCAC 780
721 ACCTACTATGCTGTAGCAGTCAAGCTGTGACAGCATCTCAGGGTCACTCTTGGCTCAC 780
781 CCCCAGGGAAGATCAGCAAGCTTATGTGACAGTGGCTGTGGCAGGAGGGGTGCT 840
781 CCCCAGGGAAGATCAGCAAGCTTATGTGACAGTGGCTGTGGCAGGAGGGGTGCT 840

841 GTGGGTACCTCGTGTACCTGATCCCTTCCGTGGCTTGCATGTTGGTCTGTGTG 900
841 GTGGGTACCTCGTGTACCTGATCCCTTCCGTGGCTTGCATGTTGGTCTGTGTG 900
901 GCTGGCTGATCTCCGTGCGGGAGCCAAAGTACCTGCGGGGTGTTGTAAACGAGTCTG 960
901 GCTGGCTGATCTCCGTGCGGGAGCCAAAGTACCTGCGGGGTGTTGTAAACGAGTCTG 960
961 GGGATTTCCACAGCTCCATCATGGGCTACAACTTCAGCTTGTGGGTCTGTTGGAG 1020
961 GGGATTTCCACAGCTCCATCATGGGCTACAACTTCAGCTTGTGGGTCTGTTGGAG 1020
1021 ATCATCTACATTTGTGCTGTGATPACCGTCGAGCGGCAATGGCATGATGGC 1080
1021 ATCATCTACATTTGTGCTGTGATPACCGTCGAGCGGCAATGGCATGATGGC 1080
1081 TTCCAGGTCCTCTCAGCATTTGGGAACTCAGCTTGGCCATCGTGTAGCTCTCAGCTCT 1140
1081 TTCCAGGTCCTCTCAGCATTTGGGAACTCAGCTTGGCCATCGTGTAGCTCTCAGCTCT 1140
1141 GGTCTCTGACAGGTTTGTCTTAAATCTTAAATATGAAAGCACCTCATGAGGCTAAA 1200
1141 GGTCTCTGACAGGTTTGTCTTAAATCTTAAATATGAAAGCACCTCATGAGGCTAAA 1200
1201 TATTTTGTGACCAAGTTTCTGAAAGTTTCTCATTGCTTGGCTTGGATTTTAA 1254
1201 TATTTTGTGACCAAGTTTCTGAAAGTTTCTCATTGCTTGGCTTGGATTTTAA 1254

RESULT 2
ADE09785

ID ADE09785 standard; DNA; 2811 BP.

XX ADE09785;

XX 29-JAN-2004 (first entry)

XX Novel DNA-related contig nucleotide sequence #507.

XX novel gene; novel protein; tissue marker; molecular weight marker;
XX chromosome marker; genetic disorder; contig; ds.

XX Unidentified.

XX WC2003054152-A2.

XX 03-JUL-2003.

XX 10-DEC-2002; 2002WO-US039555.

XX 10-DEC-2001; 2001US-0339739P.

XX 11-DEC-2001; 2001US-0339453P.

XX 14-MAR-2002; 2002US-0365091P.

XX 14-MAR-2002; 2002US-0365384P.

XX 12-APR-2002; 2002US-0372381P.

XX 22-APR-2002; 2002US-00128558.

XX 24-APR-2002; 2002US-0376045P.

XX (HYSE-) HYSEQ INC.

XX WPI; 2003-569235/53.

XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QH, Wang J;
XX Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
XX Ma Y, Wang D, Chen R, Xu C, Boyle BJ;

XX WPI; 2003-569235/53.

XX

CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responses for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
XX Sequence 487 BP; 86 A; 120 C; 156 G; 119 T; 0 U; 6 Other;

Query Match 29.2%; Score 365.8; DB 8; Length 487;
Best Local Similarity 99.5%; Pred. No. 2.7e-95;
Matches 367; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 706 AATGCCGTGTTCAACACCTACTATGCTGTAGCAGTCAGCGTGTGACAGCCATCTCAGGG 765
Db 56 ATTCCCGTGTTCACACCTACTATGCTGTAGCAGTCAGCGTGTGACAGCCATCTCAGGG 115
QY 766 TCATCCTTGCTCACCCCAAGGAGAGATCAGCAAGACTTATGTGACAGTGCAGTGTG 825
Db 116 TCATCCTTGCTCACCCCAAGGAGAGATCAGCAAGACTTATGTGACAGTGCAGTGTG 175
QY 826 GCAGGAGCGTGTGTTGGTGTACCTGTGTACCTGTGTACCTGTGTGTGCGTGTGCGCATG 885
Db 176 GCAGGAGCGTGTGTTGGTGTACCTGTGTACCTGTGTGTGCGTGTGTGCGCATG 235
QY 886 GTGCTGGTCTTGTGGTGTGGTGTACCTGTGTGTGCGTGTGTGCGTGTGTGCGCGGTGT 945
Db 236 GTGCTGGTCTTGTGGTGTGGTGTACCTGTGTGTGCGTGTGTGCGTGTGTGCGCGGTGT 295
QY 946 TGTAAACGAGTGTGCGGTATTCCTCCACAGCTTCCATCATGTGGCTACAACTTCAGCTTGTG 1005
Db 296 TGTAAACGAGTGTGCGGTATTCCTCCACAGCTTCCATCATGTGGCTACAACTTCAGCTTGTG 355
QY 1006 GGTCTGTGTGAGAGATCATCTAATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1065
Db 356 GGTCTGTGTGAGAGATCATCTAATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 415
QY 1066 AATGCCATG 1074
Db 416 AATGCCATG 424

RESULT 4
ID ABK12333 standard; cDNA; 1805 BP.
XX AC ABK12333;
XX AC ABK12333;
XX 05-JUN-2002 (first entry)
XX cDNA encoding human nonerythroid Rh glycoprotein RhBG.
XX Human; nonerythroid Rh glycoprotein; Rhbg; RhBG; antibody; immunogen;
KW Rh type B transporter gene; polytypic transporter-type protein;
KW ion transporter; chromosome 1q21.3; gene; ss.
XX
OS Homo sapiens.
XX Key Location/Qualifiers
FH 39..1415
FT CDS /*tag= a
FT /*product= "Human RhBG protein"
FT misc_feature 1248..1377
FT /*tag= b
FT /*note= "Encodes C-tail"
FT polyA_signal 1769..1774
FT /*tag= c
FT /*standard_name= "PolyA signal"

FT
XX
XX WO200216396-A2. /note= "Atypical polyadenylation site"
XX
XX 28-FEB-2002.
XX
XX 17-AUG-2001; 2001WO-US025881.
XX
XX 21-AUG-2000; 2000US-0226767P.
XX
XX (NYBL-) NEW YORK BLOOD CENT INC.
XX
XX Huang C, Liu Z;
XX
XX WPI; 2002-280905/32.
XX P-PSDB; AAU78091.
XX
XX Homologs of mouse or human nonerythroid Rh glycoproteins, Rhbg or RhBG,
XX respectively, useful in the production of antibodies which are useful for
XX detecting Rhbg or RhBG glycoproteins in a sample.
XX
XX Claim 1; Fig 1a; 59pp; English.
XX
XX The present invention relates to a new protein or peptide comprising an
XX amino acid sequence having at least 60% identity to a sequence comprising
XX 455 (mouse nonerythroid Rh glycoprotein homologue (Rhbg)) or 458 (human
XX nonerythroid Rh glycoprotein homologue (RhBG)) amino acids, fully defined
XX in the specification. The antibody of the invention is useful for
XX detecting an Rhbg or an RhBG glycoprotein in a sample, by contacting the
XX sample with antibody under conditions suitable for binding, assessing the
XX specific binding to the antibody, and thus detecting the presence of an
XX epitope of Rhbg or RhBG in the sample. The nucleic acids of the invention
XX are useful as probes for detecting transporter genes and particularly Rh
XX type B transporter genes including e.g. NH 4⁺ ion transporters. The
XX present nucleic acid sequence is that of the human RhBG gene located on
XX chromosome 1q21.3. This sequence encodes the human RhBG protein of the
XX invention. RhBG is a polytypic transporter-type protein
XX
XX Sequence 1805 BP; 351 A; 550 C; 515 G; 389 T; 0 U; 0 Other;

Query Match 14.1%; Score 177; DB 6; Length 1805;
Best Local Similarity 49.5%; Pred. No. 2.7e-40;
Matches 456; Conservative 0; Mismatches 465; Indels 0; Gaps 0;
QY 141 CTATCAAGTTGGCCCAAGATCTGACCGTGTGTCGCGCCCAATGGCTTGGCTTCCCTACCTC 200
Db 218 CTACCAAGCTTCCAGACGTCGATGCTTCTGTGGCTTGGCTTCCCTCATGGT 277
QY 201 GAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCAACCTTCTCATGTGGCGCTTGG 260
Db 278 CTTCCTGCAGCGTTACGGCTTCAGCAGCGTGGGCTTCACCTTCTCTGGCGCGCTTTGC 337
QY 261 TGTGAGTGGCAATCTCTGCTGGAGCGGCTTCTGAGCAGTTCCTCTCTGGAGAGTGGT 320
Db 338 CCTGAGTGGTCCACACTGGTCCAGGGCTTCTTCCACTCTTCCAGGTGGCCACATCCA 397
QY 321 CATCACACTGTTCAGTATTCGGCTGGCCACCATGAGTGCCTTGTGGTGTGATCTCAGT 380
Db 398 TGTGGCTGTGAGAGCATGATCATGCTGATCTTTGTGGGGGGCGGCTCATCTCCTCT 457
QY 381 GGATGCTCTCTGGGGAAGGTCAACTTGGCCAGTGTGGTGTGATGGTGTGGTGGAGGT 440
Db 458 TGGTGGCTCTCTGGGGAAGGTCAACTTGGCCAGTGTGGTGTGATGGTGTGGTGGAGGT 517
QY 441 GACAGCTTTAGCAACCTGAGGATGGTTCATCAGTAATATCTTCAACACAGACTACCAT 500
Db 518 GGTGCTGTGTGGCATCAATGATGTTGTCTCTTCTCATCTCTCTGGGGGTGAGATGCCGG 577
QY 501 GAACATGATGACATCTACTGTGTCGAGCCCTATTTTGGGCTGTGTGGCTGTGTGGCT 560
Db 578 AGGCTCCATGATATCCACACCTTTGGTGGCTTACTTTCGGGCTCGTCTCTTTCGGGTTCT 637
QY 561 GCCAAGCCTCTACCCGAGGAGACGGAGATAAAGATCAGACAGCACAGATACCAT 620

Db 638 GTACAGGCCCCAGCTGGAGAGAGACAGACCGCCAGGGCTCGCTACCAATTCAGACCT 697
 QY 621 GTCTGCCATCTGGGCGCCCTCTCTTGTGTGATGTTCTGGCCAGTTTCAACTCTGCTCT 680
 Db 698 CTTGCGCATGATTGGGACCACTCTCTGTGTGATCTCTGGCCCTAGCTTCAATGCTGCACT 757
 QY 681 GCTGAGAAGTCCAATCAAGAGGAAGATGCCGTGTTCAACACACCTACTATGCTTACAGCT 740
 Db 758 CACAGCGCTGGGGCTGGGAGCATCGAGCGGCCCTCAACACATACTACTCTCCCTGGCTGC 817
 QY 741 CAGCGTGTGACAGCCATCTCAGGGTCACTCTTGGCTCACCCCAAGGGAAGATCAGCAA 800
 Db 818 CAGCAGCCCTTGGCAGCCCTTGGCTTGTGAGCCCTTGTAGGGGAAGATGGGAGGCTTCACT 877
 QY 801 GACTTATGTGACAGTGGGCTGTGGCAGAGCGCTGTGGTGTGGTACTCTGTGTCACT 860
 Db 878 GTGCCAATCCAAATGCAGCGCTGGCTGGAGGGGTGTGTGGGACCTCAAGTGAAT 937
 QY 861 GATCCCTCTCTCGTGGCTTGCATGTTGCTGCTGTGCTGTGCTGGCTGATCTCCGTCGG 920
 Db 938 GATGCTGACACCTTGGGGCTCTGGCAGTGTGCTTGGCTGGGACTGTCTTCCAGCT 997
 QY 921 GGAGCCCAAGTACTCGCGGGGTGTGTAAACCGAGTGTGGGATTCGCCACAGTCCAT 980
 Db 998 GGGGTACAAGTCTTCAAGCCCATCTTGAATCAAAATCAAGTCCAAAGTCAACATGTGG 1057
 QY 981 CATGGGCTCAACTTCAGCTTGTGCTGTGCTGTGCTGTGAGAGATCATCTAATGTGTGCT 1040
 Db 1058 AGTCCAAACCTCCATGGGATCGCGGGGTCTGTGGGGGCCCTCTGGGGTCTCTTGTGGC 1117
 QY 1041 GGTGCTTGATACCGTCCGAGC 1061
 Db 1118 TGGACTTGCACCCATGAAGC 1138

RESULT 5
 ID ABZ11480
 XX ABZ11480 standard; cDNA; 1792 BP.
 AC ABZ11480;
 XX 20-JAN-2003 (first entry)
 DE Human polynucleotide SEQ ID NO 362.
 KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
 KW cell-proliferative disorder; neurodegenerative disease; bacterial;
 KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
 KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 KW arthritis; cytostatic; immunomodulator; neotropic; neuroprotective;
 KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
 KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
 KW antiarthritic; gene; ss.
 OS Homo sapiens.
 XX WO200270539-A2.
 PN 12-SEP-2002.
 XX 05-MAR-2002; 2002WO-US0005095.
 PF 05-MAR-2001; 2001US-00799451.
 PR (HYSE-) HYSEQ INC.
 XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehrman T, Wang J, Wang D, Drmanac RT;
 XX WPI; 2002-759812/82.
 DR P-PSDB; ABF69263.

XX PT New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
 PT or coagulation disorders.
 XX PT
 PS Claim 1; SEQ ID NO 362; 1012pp + Sequence Listing; English.
 XX CC
 CC The invention relates to an isolated polynucleotide (I) comprising a
 CC nucleotide sequence selected from any of 948 sequences (ABZ11119-
 CC ABZ12066) or their mature protein coding portion, active domain coding
 CC protein or complementary sequences. The polynucleotides are useful for
 CC identifying expressed genes or for physical mapping of human genome. The
 CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
 CC markers, as a food supplement, for generating antibodies, in medical
 CC imaging, screening and diagnostic assays and for treating cell-
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
 CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
 CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
 CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
 CC arthritis, etc. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX SQ
 SQ Sequence 1792 BP; 340 A; 547 C; 511 G; 394 T; 0 U; 0 Other;

Query Match 14.0%; Score 175.4; DB 6; Length 1792;
 Best Local Similarity 49.4%; Pred No 8e-40;
 Matches 455; Conservative 0; Mismatches 466; Indels 0; Gaps 0;
 QY 141 CTATCAAGTTGGCCAGATCTGACCGTGATGGCGCCATTTGGCTTCTCTCACCTC 200
 Db 211 CTACCAAGCTTCCAGAGCTGCATGCCATGGTCTTGTGGGCTTTGACTTCTCATGT 270
 QY 201 GAGTTTCCGGACACAGCTGAGCAGTGTGGCTTCAACCTTCTCATGCTGGCGCTGG 260
 Db 271 CTTCTGACGCTTACGGCTTCAGCAGCTGGGCTTCACTTCTTCCAGGTGGCCATCA 330
 QY 261 TGTGAGTGGGCAATCTCTGTGGAGCGCTTCTGAGCCAGTTCCTTCTGGGAAGTGT 320
 Db 331 CCTGCAGTGTCCACACTGGTCCAGGCTTCTTCACTTCTTCCAGGTGGCCATCA 390
 QY 321 CATCACATGTTCAATGTTGGCTGGCCACCATGAGTCTTGTCTGGTCTGATCTCAGT 380
 Db 391 TGTTCGCTGGAGAGCATGATCAATGCTGACTTTTGTGGGGGCGCTCATCTCCTTT 450
 QY 381 GGATCTCTCTTGGGGAAGGTCACTTGGCGAGTGTGGTGTGATGTTGGTGGAGGT 440
 Db 451 TGGTCCCGCTCTGGGCAAGACCGGGCCCTACCCAGCTGTCTCATGCGCCCTGCTGGAGT 510
 QY 441 GACAGCTTTAGGCAACCTTGAGGATGCTCATCAGTAATATCTTCAACACAGACTACCA 500
 Db 511 GGTGCTGTTGGCATCAATGATTTGTCTCTTCACTTCTCTGGGGTGAGAGTCCCG 570
 QY 501 GAACATGATGACATCTACGTTGTCGAGCCCTATTTTGGGCTGTCTGTGGCTGTGCT 560
 Db 571 AGGCTCCATGACTATCCACACTTTTGGTGCCTTCTTGGGCTCGTCTTTCGCGGGTCT 630
 QY 561 GCCAAAGCCTCTACCCGAGGAGCGGAGGATTAAGATCAGACAGCAAGATACCACTTT 620
 Db 631 GTACAGGCCCGCTGGGAGAGCAAGCACCGCCAGGGCTCCGCTTACCAATCAGACCT 690
 QY 621 GTCTGCCATGCTGGGCGCCCTCTTCTTGTGGATGTTCTGGCCAAAGTTTCAACTCTGCT 680
 Db 691 CTTGCCATGATTGGGACCATCTTCTGTGGATCTTCTGGCTAGCTTCAATGCTGACT 750
 QY 681 GCTGAGAAGTCCAAATCGAAGGAAGATGCGGTGTTCAACACCTTACTATGCTGTAGCAGT 740
 Db 751 CACAGCGCTGGGGCTGGGAGCATCGGACCGCCCTCAACACATACTTCTCTGGCTGC 810
 QY 741 CAGCGTGTGACAGCCATCTCAGGGTCTATCTTGGCTCACCCCCAAGGAAGATCAGCAA 800

Db 811 CAGCACCCTTGGACACCTTTGCCTTGTTCACCCCTTTAGGGAGATGGGAGCTTGACAT 870
QY 801 GACTTATGTGCAGTGCCTGTGTTGCGCAGAGCGTGGCTGGGTAACCTGTGTCACT 860
Db 871 GGTCCACATCCAAATGACGCGCTGGCTGGAGGGTGTGGTGGGACCTCAAGTGAAT 930
QY 861 GATCCCTTCTCGTGGCTTGCATGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 920
Db 931 GATGCTGACACCTTTGGGGCTCTGGCAGCTGGCTTCTTGGCTGGGACTGTCTCCAGCT 990
QY 921 GGGAGCAAGTACTCTCGCGGGTGTGTAAACCGAGTGTGGGATTCCTCCACAGCTCCAT 980
Db 991 GGGGTACAGTCTTTCACGCCCATCTTGAATCAAAATTCAGAGTCCAGACACATGTGG 1050
QY 981 CATGGGCTCAACTCAGCTTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1040
Db 1051 AGTCCACACCTCCATGGATGCGGGGTCTCTGGGGGCGCTCTGGGGTCTCTGGTGGC 1110
QY 1041 GTGCTGTGATACCTGGAGC 1061
Db 1111 TGGACTTGGCACCACCAATGAAGC 1131

RESULT 6
ACD20392/c
ID ACD20392 standard; DNA; 1351 BP.
XX AC ACD20392;
XX XX
DT 26-AUG-2003 (first entry)
XX
DE DNA encoding human NOV15 protein.
XX
KW Human; NOVX; inflammatory disorder; demyelination disease; stroke;
KW renal disorder; infection; cardiomyopathy; atherosclerosis; acne;
KW hypertension; pancreatitis; Von Hippel-Lindau; endometriosis; fertility;
KW scleroderma; cirrhosis; inflammatory bowel disease; Crohn's disease;
KW haemophilia; autoimmune disease; allergy; AIDS;
KW Graft versus host disease; Alzheimer's disease; arthritis; pain;
KW Parkinson's disease; Huntington's disease; obesity; diabetes;
KW hair growth; hair loss; asthma; schizophrenia; glomerulonephritis;
KW lupus erythematosus; psoriasis; antidiabetic; anorectic; metabolic;
KW neutropenic; neuroprotective; cytostatic; antibacterial; virucide;
KW proteoacid; antiarteriosclerotic; hypotensive; cerebroprotective;
KW antiinflammatory; gynaecological; antiinfertility; dermatological;
KW hepatotropic; haemostatic; immunosuppressive; antiallergic;
KW antiarthritic; anticonvulsant; antiseborrhoeic; antiasthmatic;
KW neuroleptic; anti-HIV; analgesic; nephrotropic; antipsoriatic; gene; ds.
XX
OS Homo sapiens.
XX
XX WO200298917-A2.
XX
PD 12-DEC-2002.
XX
PF 12-FEB-2002; 2002WO-US022049.
XX
PR 12-FEB-2001; 2001US-0268221P.
PR 13-FEB-2001; 2001US-0268496P.
PR 14-FEB-2001; 2001US-0268646P.
PR 14-FEB-2001; 2001US-0268665P.
PR 15-FEB-2001; 2001US-0269136P.
PR 16-FEB-2001; 2001US-0269310P.
PR 16-FEB-2001; 2001US-0269530P.
PR 15-MAR-2001; 2001US-0276403P.
PR 15-MAR-2001; 2001US-0276399P.
PR 16-MAR-2001; 2001US-0276703P.
PR 23-MAR-2001; 2001US-0278199P.
PR 28-MAR-2001; 2001US-0279274P.
PR 30-MAR-2001; 2001US-0280238P.
PR 02-APR-2001; 2001US-0280899P.
PR 08-AUG-2001; 2001US-0310797P.
PR 14-AUG-2001; 2001US-0312284P.

PR 14-SEP-2001; 2001US-0322294P.
PR 14-SEP-2001; 2001US-0322295P.
PR 18-OCT-2001; 2001US-0330293P.
PR 31-OCT-2001; 2001US-033104P.
PR 31-OCT-2001; 2001US-033109P.
PR 21-NOV-2001; 2001US-033127P.
PR 28-NOV-2001; 2001US-0331772P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Guo X, Fernandes E, Li L, Kekuda R, Liu Y, Leite M, Spytek KA;
PI Ji W, Casman SJ, Boldog FL, Patturajan M, Vernet CM, Ballinger RA;
PI Malyankar UM, Tchernev VT, Bialock AD, Gusev VI, Rastelli L;
PI Mezes PD, Ellerman K, Heyes M, Herrmann JL, Shinkets RA, Ioime N;
PI Pena CEA, Shenoy SG, Taupier RJ, Gerlach V, Gorman L;
XX
XX WPI; 2003-148650/14.
DR P-PSDB; ABO15004.
XX
XX Novel NOVX polypeptide useful for identifying an agent that binds to the
PT polypeptide, and for treating cardiomyopathy, atherosclerosis,
PT hypertension, infertility, scleroderma, cirrhosis, and inflammatory bowel
PT disease.
XX
XX Claim 9; Page 115; 566pp; English.
XX
XX The present invention relates to the isolation of novel human
CC polypeptides referred to as NOVX (NOV1-NOV37), variants of these
CC proteins, and the polynucleotide sequences encoding them. The NOVX
CC proteins of the invention share homology to various types of protein
CC families such as zinc finger-like proteins, enzymes, receptors, and
CC lipoproteins. The sequences of the invention may be useful in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease. For example they can be used to treat inflammatory
CC disorders, demyelination disease, renal disorders, infections,
CC cardiomyopathy, atherosclerosis, hypertension, stroke, pancreatitis, Von
CC Hippel-Lindau, endometriosis, fertility, scleroderma, cirrhosis,
CC inflammatory bowel disease, Crohn's disease, haemophilia, autoimmune
CC diseases, allergies, graft versus host disease, Alzheimer's disease,
CC arthritis, Parkinson's disease, Huntington's disease, obesity, diabetes,
CC acne, hair growth/loss, asthma, schizophrenia, AIDS, pain,
CC glomerulonephritis, lupus erythematosus, and psoriasis. ACD20372-ACD20427
CC represent DNA sequences encoding the NOVX polypeptides of the invention.
CC Note: SEQ ID Nos 113-460 are known sequences used for homology purposes
XX
XX Sequence 1351 BP; 325 A; 372 C; 389 G; 265 T; 0 U; 0 Other;
XX
Query Match 13.2%; Score 165.4; DB 7; Length 1351;
Best Local Similarity 49.1%; Pred. No. 5.4e-37;
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;
QY 141 CTATCAAGTTGGCCAGATCTGACCGTGTGGCGCCATTGGCTTGGCTTCTCTCACTC 200
Db 1175 CTACCAAGCTTCCAGACGTGCACGTGATGTTCTCGTGGCTTGGCTTCTCTCACTC 1116
QY 201 GAGTTTCCGGAGACACAGCTGGAGCAGTGGCCCTTCACTCTTCATGCTGGCTTGG 260
Db 1115 TTTCTCGAGCGCTACGGCTTCAGCGCGCTTCACTTCTGTGGCAGCTTGG 1056
QY 261 TGTGACGTGGCAATCTGCTGGAGCGCTTCTGAGCAGTTCCTTCTGGGAAGTGGT 320
Db 1055 CATCAGTGGCGCTGCTCATGACGGCTGGTTTCACTTCTTAAGACCGCTACATGCT 996
QY 321 CATCACACTGTTTCACTATTCGGCTGGCCACCATGAGTCTTGTGGTGTGATCTCAGT 380
Db 995 CGTGGCGTGGAGAACCTCATCAACGCTGACTTCTGCGTGGCTTCTGTCTGGTGGCTT 936
QY 381 GGATGCTCTCTGGGGAAGGTCACTGGCCAGTGTGGTGTGATGCTGGTGGTGGT 440
Db 935 TGGGCGAGTCTTGGGTAAGTCAAGCCCAATTCAGCTGCTCATGATCTTCTTCAAGT 876
QY 441 GACAGCTTTAGCAACCTGAGGTGGTTCATCAGTAATATCTTCAACACAGATACCAT 500

Db 875 GACCCTCTCGTGTGAATGAGTTCAATTCCTTAACTGCTAAAGGTGAAGATGACG 816
Qy 501 GACATGATGACATCTACATGTTTCGACGCTATTTGGCTGTCTGTGGCTTGGTGCCT 560
Db 815 AGGCTCATGACCATCCACATTTTGGCGCTACTTTTGGCTCACAGTGACCGGATCCT 756
Qy 561 GCCAAAGCTCTTACCCAGGGAACGGAGGATAAAGATCAGACAGCAACGATACCCAGTTT 620
Db 755 CTACCGAGCAACTAGAGCAGAGCAGAGGAGACAGAAATTCGTGTACCGATCGGACCT 696
Qy 621 GTCTGCATGCTGGGCGCCTCTCTTGTGTGATGTTCTGTGCGCAAGTTTCAACTCTGTCT 690
Db 695 CTTTGGCATGATTTGGCACCCTCTCTCTGTGATGTAATGCGCCAGCTTCAACTCAGCCAT 636
Qy 681 GCTGAGAAGTCCNAATCAAGAGGAGATGCCGTGTTCAACACTACTATGCTGTAGAGT 740
Db 635 ATCCTACCTATGGGACAGCCAGCAGCCAGCGCCCATCAACACTACTGCTCTTGGCAGC 576
Qy 741 CAGCGTGTGACAGCCATCTCAGGCTCATCCTTTGGCTCACCCCAAGGGAAGATCAGCAA 800
Db 575 CTGCGTGTACTCTCGTGGCATATCCAGTGCCTGCAAGAGAGGCAAGCTGGACAT 516
Qy 801 GACTTATGTGCACAGTGGGTGTGGCAGAGCGGTGGCTGTGGGTACCTCTGTGTACCT 860
Db 515 GGTGCATATCCAGATGCCACGCTCGCAGAGGGGTGGCGTGGGTACCGTGTGAGAT 456
Qy 861 GATCCCTTCTCGTGGCTTGCATGCTGTGGTCTTGTGGTGTGGCTGTATCTCGTGG 920
Db 455 GATGCTCATGCTTACGGTGGCCCTCATCCTGCGCTTCTGTGCGGATCATCTCCACCT 396
Qy 921 GGGAGCAAGTACTCGCGGGGTGTGTAAACCGAGTGTGGGATTCGCCACAGCTCCAT 980
Db 395 GGGTTTGTATACCTGACCCCATCTCTGAGTCCCGCTGCATCCAGACACATGTGG 336
Qy 981 CATGGCTACAATTGAGTGTGGGTGTGGTGTGGTGTGGAGATCATCTACATTGTG 1035
Db 335 CATTACAATCTGATGCGATTCCTGGCATCATAGGCGGATCGTGGGTGCTGTG 281

RESULT 7
AAC64286
ID AAC64286 standard; cDNA; 1853 BP.
AC AAC64286;
XX
DT 23-FEB-2001 (first entry)
XX
DE Human membrane-associated protein HUMAP-13 cDNA.
XX
KW Human membrane-associated protein; HUMAP; transgenic organism;
KW drug screening; cell signalling modulator; agonist; antagonist;
KW cell differentiation modulator; cell proliferation modulator;
KW cell proliferative disorder; cancer; cell differentiation disorder;
KW developmental disorder; cell signalling disorders; endocrine disorder;
KW hyperpituitarism; hypothyroidism; hyperparathyroidism; infection;
KW pancreatic disorder; diabetes mellitus; immunological disorder;
KW hereditary neuropathy; gonadal steroid hormone associated disorder;
KW infertility; ss.
XX
OS Homo sapiens.
XX
PN WO200065054-A2.
XX
PD 02-NOV-2000.
XX
PF 20-APR-2000; 2000MO-US010884.
XX
PR 23-APR-1999; 99US-0130694P.
PR 23-JUN-1999; 99US-0140580P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Hillman JL, Bandman O, Tang YT, Lal P, Yue H, Reddy R;

PI Azimzai Y, Baughn MR;
XX WPI; 2000-687346/67.
DR P-PSDB; AAB29656.
XX
PT Human membrane-associated protein, useful for diagnosis and treatment of
PT cell signaling, cell differentiation and cell proliferation disorders
PT such as cancer, and for identifying agonists and antagonists.
XX
PS Claim 4; Page 96; 99pp; English.
XX
XX The invention relates to 17 human membrane-associated proteins, HUMAP-1
CC to HUMAP-17 (AAB29644-B29660) and the cDNAs encoding them (AAC64274-
CC C64290). The invention also relates to expression constructs, host cells
CC and transgenic organisms comprising a HUMAP nucleic acid sequence; the
CC recombinant preparation of a HUMAP; methods of screening compounds for
CC their ability to modulate HUMAP activity or expression; and
CC pharmaceutical compositions comprising a HUMAP protein, a HUMAP agonist
CC or HUMAP antagonist. The HUMAPs acts as modulators of cell signalling,
CC differentiation and proliferation. A HUMAP is useful for screening a
CC compound for effectiveness as an agonist or antagonist of HUMAP activity.
CC The protein, or the identified agonist or antagonist is useful for
CC treating a disease or condition associated with decreased or increased
CC expression of functional HUMAP. A HUMAP nucleic acid is useful for
CC screening a compound for its ability to alter expression of that
CC particular HUMAP gene. A wide variety of disease may be treated using
CC compositions of the invention. These diseases include cell proliferative
CC disorders (e.g., actinic keratosis, arteriosclerosis); cancer (e.g.,
CC breast, bladder, bone marrow, brain and uterine cancer); cell
CC differentiation disorders, in particular developmental disorders (e.g.,
CC renal tubular acidosis, anaemia, Cushing's syndrome, achondroplasia,
CC epilepsy, and muscular dystrophy); cell signalling disorders, in
CC particular endocrine disorders such as hypothalamus and pituitary
CC disorders resulting from lesions such as thrombosis; disorders associated
CC with hyperpituitarism (e.g., acromegaly); disorders associated with
CC hypothyroidism (e.g., goitre); hyperparathyroidism; pancreatic disorders
CC such as type I or type II diabetes mellitus; infections; immunological
CC disorders; hereditary neuropathies (e.g., neurofibromatosis); and
CC disorders associated with gonadal steroid hormones (e.g., infertility,
CC endometriosis, polycystic ovary syndrome, osteoporosis, Leydig cell
CC deficiency and gynecomastia). Antibodies which specifically bind HUMAP
CC may be used for the diagnosis of disorders associated with the expression
CC of HUMAP, or in assays to monitor patients being treated with HUMAP or
CC agonists, antagonists or inhibitors of HUMAP. The present sequence
XX represents a HUMAP cDNA of the invention
SQ
SQ Sequence 1853 BP; 367 A; 563 C; 508 G; 415 T; 0 U; 0 Other;
Query Match 13.2%; Score 165.4; DB 3; Length 1853;
Best Local Similarity 49.1%; Pred. No. 6.4e-37;
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;
Qy 141 CTATCAAGTTGGCCCAAGATCTGACCGTGTATGCGCGCCATTGGCTTGGCTTCCCTCACCTC 200
Db 207 CTACCAAGCTTCAGGACGTGCATGATGCTTCTGTTGGCTTGGCTTCCCTCATGAC 266
Qy 201 GAGTTTCCGAGACACAGCTGGAGCAGTGTGGCTTCAACCTTTCATGCTGGCGCTTGG 260
Db 267 TTTCTCTGAGCGCTACGGCTTCAGCGCGTGGCTTCAACTTTCGTTGGCAGCTTCGG 326
Qy 261 TGTGAGTGGGCAATCTCTGTGGAGCGCTTCTGAGCAGTTCCTTCTGGGAAGGTGGT 320
Db 327 CATCCAGTGGGCGCTGCTCATGACGGGCTGGTTCCACTTCTTACAAGACCGTACATCGT 386
Qy 321 CATCACACTGTTTCACTATTTCGGCTGGCCACCATGAGTGTCTTGTGGTGGTGTCTCAGT 380
Db 387 CGTGGCGTGGAGAACCTCATCAACGTGACTTCTGCGTGGCGCTCTGTCTCGTGGCGCTT 446
Qy 381 GGATGCTCTCTTGGGGAAGGTCAACTTGGCGCAGTTGGTGGTGTGGTGTGGTGGAGGT 440
Db 447 TGGGCGAGTTCTGGGTAAGTCAGCCCATTCAGCTGCTCATCATGACTTCTTCCAGT 506
Qy 441 GACAGCTTTAGCAACCTGAGGATGGTTCATCAGTAATATCTTCAACACAGACTACCAT 500

QY 981 CATGGCTACAACTTCAGTTCCTGGTCTGCTGGAGAGATCATCTAATTGTG 1035
 Db 1057 CATTAAACATCTGCATTCCTGGCATCATAGCGGCATCGTGGGTGCTGTG 1111

RESULT 9

ABK49215
 ID ABK49215 standard; cDNA; 1952 BP.

XX AC ABK49215;

XX AC
 XX DT 15-JUL-2002 (first entry)

XX DE cDNA encoding human Rh type C gene (RHCG) protein.

XX RHCG: human; non-erythroid Rh type C glycoprotein; chromosome 15q25;
 KW gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT CDS 24..1440

FT /*tag= a

FT /product= "RhCG protein"

FT polyA_signal 1985..1990

FT /*tag= b

XX WO200220719-A2.

XX PN 14-MAR-2002.

XX PD 05-SEP-2001; 2001WO-US027503.

XX PF 07-SEP-2000; 2000US-0230660P.

XX PR (NYBL-) NEW YORK BLOOD CENT INC.

XX PA Huang C, Liu Z;

XX PI WPI; 2002-351774/38.

XX DR P-PSDB; AAU78997.

XX Nucleic acid sequences encoding novel mammalian nonerythroid Rh type C
 PT and glycoproteins which have a characteristic twelve transmembrane domain
 PT structure.

XX Claim 2; Fig 1; 53pp; English.

XX This invention relates to the nucleic acid and protein sequences of novel
 CC human and mouse non-erythroid Rh type C glycoprotein (RhCG). The RhCG
 CC protein and the mouse homologue (rhcg) have a characteristic 12
 CC transmembrane domain structure and are expressed in kidneys and testis.
 CC The invention also comprises a method for antibody that specifically
 CC binds an epitope of the glycoprotein and a method for detecting the
 CC protein using this antibody. The antibodies of the invention may be used
 CC in Western blots, enzyme linked immunosorbent assays (ELISA) or
 CC immunohistochemical assays to identify the non-erythroid tissues,
 CC particularly kidney and testis, that express the RhCG or Rhcg
 CC glycoprotein. The methods are used for detecting an Rhcg or and RhCG
 CC glycoprotein in a sample. The present sequence represents the cDNA
 CC encoding the human RhCG Rh type C glycoprotein (RhCG) protein sequence of
 CC the invention. The gene encoding this protein is located on human
 CC chromosome 15q25

XX SQ Sequence 1952 BP; 398 A; 586 C; 524 G; 444 T; 0 U; 0 Other;

XX Query Match

XX Best Local Similarity 13.2%; Score 165.4; DB 6; Length 1952;

XX Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGCCAAGATCTGACCGTGATGGCGGCATCTGGGTCTCTCACCTC 200
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 201 CTATCCAAAGCTTCAGGAGGTGCACGTGATGTCTTCTGGGGCTTCGGCTTCTCATGAC 260
 QY 201 GAGTTTCGGGAGACACAGCTGGAGCAGTGTGCCCTTCAACCTCTTTCATGCTGCGCTGG 260
 Db 261 TTCTCTGCAGCGCTACGGCTTCAGCGCGCTGGGCTTCACTTCTCTGTGGCAGCTTCGG 320
 QY 261 TGTGCACTGGGCAATCTCTGCTGGAGGGCTTCTGAGCCAGTTCCTTCTGGGAAAGTGTG 320
 Db 321 CATCCAGTGGGCGCTGCTCATGACGGGTGGTTCACATTCCTTACAGACCGCTACATCTG 380
 QY 321 CATCACACTGTTCACTATTTCGGCTGCGCACCATGATGCTGTTGTGGTCTGATCTCACT 380
 Db 381 CGTGGCGGTGGAGAACCTCATCAAGCTGACTTCTGCGTGGCTCTGTCTGGTGGCTT 440
 QY 381 GGATCTGTCTTGGGAAAGGTCAACTTGGCGCAGTTGTTGGTGTGATGGTCTGGTGGAGGT 440
 Db 441 TGGGCGAGTCTGGGTAAAGTCAGCCCCATTTCAGCTGCTCATCATGACTTTCTTCCAAGT 500
 QY 441 GACAGCTTTAGCAACCTGAGGATGGTTCATCATGTAATATCTTCAACAGACTACCAT 500
 Db 501 GACCTCTTCGCTGTGAATGATTCATTCTCTTAACCTGTAAAGGTGAAGGATGCAGG 560
 QY 501 GAAATGATGACATCTACGTGTTTCGACGCTTATTTTGGGCTGTCTGTGGCTGTGCT 560
 Db 561 AGGCTCCATGACCATCCACATTTTGGGCTTCTTTGGGCTTCAAGTACCCGATCTCT 620
 QY 561 GCCAAAGCTCTTACCCGAGGGAACGAGGATAAAGATCAGACAGCAAGATACCCAGTTT 620
 Db 621 CTACCGACGCAACCTTAGAGCAGCAGCAGAGACAGAAATTCGTGTACCACTGGACCT 680
 QY 621 GTCTGCCATGTGGGGGCGCTCTTCTGTGATGTTCTGGCCAAAGTTTCAACTCTGCTCT 680
 Db 681 CTTTGGCATGATTTGGCACCCCTCTTCTGTGGATGACTGGCCACGCTTCACTCAGCCAT 740
 QY 681 GCTGAGAAAGTCCAAATCGAAAGGAAGATGCGGTGTTTCAACACTTACTATGCTGTAGCAGT 740
 Db 741 ATCTTACCATGGGACAGCCAGCAGCCGACATCACCTACTGCTCTTGGCAGC 800
 QY 741 CAGCGTGTGACAGCCATCTTCAGGGTCACTCTTGGCTCACCCCCAAGGAAAGATCAGCAA 800
 Db 801 CTGCGTGTCTTACCTCGGTGGCAATATCCAGTCCCTGCACAAAGAGGCAAGCTGGACAT 860
 QY 801 GACTTATGTGCACAGTCGGTGTGGCAGAGGCGCTGTGGGTACTCTGTGTCACCT 860
 Db 861 GGTGCATATCCAGATGCCAGCTCGCAGAGGGGTGGCGGTGCGGTGCTGCTGAGAT 920
 QY 861 GATCCCTTCTCGGTGGCTTGCCATGGTGTGCTGTGGCTGGGCTGATCTCGTCTGG 920
 Db 921 GATGCTCATGCTTACGGTGGCTTCACTCATCGGCTTCTGTGGGCATCATCTCCACCT 980
 QY 921 GGGAGCCCAAGTACCTGCGGGGTGTTGTAACCGAGTGTCTGGGGATTCCTCCACAGTCCAT 980
 Db 981 GGGTTTGTATACCTGACCCCATTCCTGGAGTCCCGGTGCACATCCAGGACACATGTGG 1040
 QY 981 CATGGCTTACAACCTTCAGCTGCTGGTCTGTTGGAGAGATCATCTACATGTG 1035
 Db 1041 CATTAAACAATCTGCATGCAATTCCTGGCATCATAGGCGGCATCTGTTGGTGTGCTG 1095

RESULT 10

ABX76257

ID ABX76257 standard; DNA; 1968 BP.

XX AC ABX76257;

XX DT 02-APR-2003 (first entry)

XX Lung cancer-associated polynucleotide #123.

XX Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 KW antinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;

XX PD 05-OCT-2000.
XX PF 31-MAR-2000; 2000WO-US008621.
XX PR 31-MAR-1999; 99US-0127607P.
XX PR 02-APR-1999; 99US-0127636P.
XX PR 05-APR-1999; 99US-0127728P.
XX PR 30-MAR-2000; 2000US-00540763.
XX PR (CURA-) CURAGEN CORP.
XX PA Shimkets RA, Leach M;
XX PI WPI: 2000-602362/57.
XX DR P-PSDB; AAB42135.
XX DR Novel nucleic acids and peptides derived from open reading frame X,
XX PT useful for treating e.g. cancers, proliferative disorders,
XX PT neurodegenerative disorders and cardiovascular disease.
XX PS Claim 5; Page 2946-2948; 5507pp; English.
XX CC AAC/4446 to AAC/77606 encode the proteins given in AAB40237 to AAB43397,
XX CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX CC sequences have activities such as: cytosolic; hepatotropic; vulnery;
XX CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
XX CC anticonvulsant; antiarthritic; immunosuppressive; antidiabetic; hypotensive;
XX CC cardiac; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
XX CC dermatological; immunosuppressive; antirheumatic; antithyroid; and antianemic. The
XX CC sequences can be used for determining the presence of or predisposition
XX CC to, or preventing or treating pathological conditions associated with an
XX CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
XX CC proteins in gene therapy vectors. The proteins and nucleic acids may be
XX CC used to treat cancers, proliferative disorders, neurodegenerative
XX CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
XX CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
XX CC storage, systemic lupus erythematosus, severe combined immunodeficiency
XX CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
XX CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
XX CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
XX CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX SQ Sequence 1970 BP; 400 A; 593 C; 530 G; 445 T; 0 U; 2 Other;

Query Match 13.2%; Score 165.4; DB 3; Length 1970;
Best Local Similarity 49.1%; Pred. No. 6.e-37;
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGGCCAGATCTGACCGTGTATGGCGCCATTCGCTGGGCTTCCTCACCTC 200
Db 219 CTACCAAGTTCAGGACGTGACGATGATGCTTCGCGGCTTCGCTTCCTCATGAC 278
QY 201 GAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCACCTTCATGCTGGGCTTGG 260
Db 279 TTCTCTGACGCGTACGCTTCAGGCGCGTGGCTTCACCTTCCTGGGAGCCTTCGG 338
QY 261 TGTGAGTGGGCAATCTGCTGGAGCGTTCCTGAGCCAGTTCCTCTGGGAAGGTGGT 320
Db 339 CATCCAGTGGGCGCTGCTCATGACGGCTGGTTCACCTTCATCAAGACCGCTACATGCT 398
QY 321 CATCACATGTTTCAGTATTCGGCTGGCCACCATGATGCTTGTTCGGTGTGATCTCAGT 380
Db 399 CGTGGGCGTGGAGAACCTTCACGCTGACTTTCGCGTGGCCCTCTGCTGGCTGGCCCTT 458
QY 381 GGATGCTCTCTTGGGGAAGGTCAAATTGGCGCAGTGTGGTGTGATGCTGGTGGAGGT 440
Db 459 TGGGSCAGTTCTGGGTAAGTACGCCCCATTCAGCTGCTCATCATGACTTCTTCCAGT 518
QY 441 GACAGCTTTAGGCACTCGAGATGGTTCATCAGTAATATCTTCAACACAGACTACCAT 500
Db 519 GACCTCTCTGCTGATGAGTTCATCTCTTAACTGCTAAAGGTGAAGGATGCAGG 578

QY 501 GAACATGATGCATCTACGTGTTTCGACGCTATTTTGGCTCTCTGTGGCTGGTGCCT 560
Db 579 AGGCTCCATGACCATCCACATTTGGCGCTACTTTGGGCTCACAGTGAACCGGATCCT 638
QY 561 GCCAAGCCTCTTACCCGAGGGAACGAGGATAAAGATCAGACAGCAACATACCCAGTTT 620
Db 639 CTACCCAGCAACCTTAGACGAGCAAGGAGAGACAGAAATCTGTGTACCCAGTCGGACCT 698
QY 621 GTCTGCCATGCTGGGCGCCCTCTTCTGTGGATGTTCTGGCCAAAGTTTCAACTCTGCTCT 680
Db 699 CTTTGGCATGATGGCACCTCTTCTGTGGATGTTCTGGCCAAAGTTTCAACTCTGCTCT 758
QY 681 GCTGAGAACTCCAATCGAAAGGAAGATGCCGTGTTTCAACACTACTATGCTGTAGCACT 740
Db 759 ATCTTACCATGGGACAGCCAGCAGCCAGCGCCCATCAACACTACTGCTCTTGGCAGC 818
QY 741 CAGCGTGTGACAGCATCTCAGGGTTCATCTTGGCTCAACCCCAAGGAAATCAGCAA 800
Db 819 CTGCGTGTCTTACCTCGGTGGCAATATCCAGTCCCTGTCACAAAGAGGCAAGCTGGACAT 878
QY 801 GACTTATGTGCACAGTGGGTGTTGGCAGGAGCGTGGTGGTGGTGGTGGTGGTGGTGG 860
Db 879 GGTGCACATCCAGAAATGCCACCTCGCAGGAGGGGTGGCGGTGGGTGGTGGTGGTGG 938
QY 861 GATCCCTTTCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 920
Db 939 GATGCTCATGCTTACGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 998
QY 921 GGGAGCCCAAGTACCTGCGGGGTGTTTAAACGAGTGTGGGGATTCGCCAGAGTCCCAT 980
Db 999 GGGTTTGTATACCTGAGCCCATTCCTGGAGTCCCGGTGCACATCCAGGACACATGTGG 1058
QY 981 CATGGCTTACAACTTCACTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1035
Db 1059 CATTAACAATTCGATGCGATTCCTGGCATCATAGGCGGCTTCGTGGGTGGTGGTGG 1113

RESULT 12
ABK12334
ID ABK12334 standard; cDNA; 1679 BP.
XX AC ABK12334;
XX XX
XX DT 05-JUN-2002 (first entry)
XX DE cDNA encoding mouse nonerythroid Rh glycoprotein Rhbg.
XX KW Mouse; nonerythroid Rh glycoprotein; Rhbg; RhBG; antibody; immunogen;
XX KW Rh type B transporter gene; polytypic transporter-type protein;
XX KW ion transporter; chromosome 3; gene; ss.
XX OS Mus sp.
XX PH Key Location/Qualifiers
XX FT CDS 39..1531
XX FT /*tag= a
XX FT /product= "Mouse Rhbg protein"
XX FT polyA_signal 1634..1639
XX FT /*tag= b
XX FT /standard_name= "polyA signal"
XX FT /note= "Atypical polyadenylation site"
XX XX WO200216396-A2.
XX PN 28-FEB-2002.
XX PD 17-AUG-2001; 2001WO-US025881.
XX PF 21-AUG-2000; 2000US-0226767P.
XX PR (NYBL-) NEW YORK BLOOD CENT INC.
XX PA
XX XX

PI Huang C, Liu Z;
XX WPI: 2002-280905/32.
DR P-PSDB; AAU78092.
XX
PT Homologs of mouse or human nonerythroid Rh glycoproteins, Rbpg or Rbpg.
PT respectively, useful in the production of antibodies which are useful for
PT detecting Rbpg or Rbpg glycoproteins in a sample.
XX
PS Claim 1; Fig 1a; 59pp; English.
XX
CC The present invention relates to a new protein or peptide comprising an
CC amino acid sequence having at least 60% identity to a sequence comprising
CC 455 (mouse nonerythroid Rh glycoprotein homologue (Rbpg)) or 458 (human
CC nonerythroid Rh glycoprotein homologue (Rbpg)) amino acids, fully defined
CC in the specification. The antibody of the invention is useful for
CC detecting an Rbpg or an Rbpg glycoprotein in a sample, by contacting the
CC sample with antibody under conditions suitable for binding, assessing the
CC specific binding to the antibody, and thus detecting the presence of an
CC epitope of Rbpg or Rbpg in the sample. The nucleic acids of the invention
CC are useful as probes for detecting transporter genes and particularly Rh
CC type B transporter genes including e.g. NH 4⁺ ion transporters. The
CC present nucleic acid sequence is that of the mouse Rbpg gene located on
CC chromosome 3. This sequence encodes the mouse Rbpg protein of the
CC invention. Rbpg is a polytypic transporter-type protein
XX
SQ Sequence 1679 BP; 324 A; 491 C; 469 G; 395 T; 0 U; 0 Other;
Query Match 12.7%; Score 159.4; DB 6; Length 1679;
Best Local Similarity 48.3%; Pred. No. 3.3e-35;
Matches 445; Conservative 0; Mismatches 476; Indels 0; Gaps 0;
141 CTATCAAGTTGGCCAAAGATGACCGGTGATGGGGCCCATTTGGCTTGGCTTCTCACCTC 200
209 CTACCAAGCTTCCAGGATGACGCTGATGCTTCTGGGGCTTGGCTTCTCACCTC 268
201 GAGTTTCCGAGACACGCTGGAGCTGGCTTCAACCTCTTCATGCTGGCGCTTGG 260
269 CTTTCTACAGCGGTACGGCTTACGAGTGGGGTTTACCTTCTGCTGGCCAGCTCAC 328
261 TGTGCGAGTGGCAATCTGCTGGACGGCTTCTTCCAGGCTTCTTCTGGGAAGTGGT 320
329 CTTGCGAGTGGCCACACTGTCTCAAGGCTTCTTCCAGCTTCCAGGCTGACATCCA 388
321 CATCACACTGTTAGTATTGGCTGGCCACCATGAGTCTTGTGCTGCTGATCTCAGT 380
389 TGTGCGGTGGAGAGTTTGTATCAACGCTGACTTCTGCGGGGAGCTGTGCTCATCTTT 448
381 GGATGCTGCTTGGGGAGGTCAACTTGGGCGAGTGGTGGTGGTGGTGGTGGGAGGT 440
449 CGGGGCTGTTCTGGGCAAGACTGGGCGAGCCAGCTGCTGCTAATGCTCTACTGGAGC 508
441 GACAGCTTTAGGCAACCTGAGGATGGTATCAGTAATATCTTCAACACAGACTACCAT 500
509 AGTGTGTTAGGTCAACGAGTTTATATCTACTCAGTCTCTCTGGGGGTGAGAGATGCT 568
501 GAAATGATGACATCTACGTTTGGCAGCTTATTTGGGCTGCTGTGGCTGCTGCTGCT 560
569 AGGTGCTCAATCAATCACATTTGGGGCTTCTTGGGGCTGCTTCTTCCAGGCTCT 628
561 GCCAAGGCTCTACCCAGGGAAGGAGGATAAAGATCAGACAGCAACGATACCCAGTTT 620
629 CTACAGATCCAGCTGGAGAGAGCAGCAGCATCGCAGAGCTCTGTCTACAACTCTGAC 688
621 GTCTGCCATGTGGGCGCCCTCTTCTTGTGGAGTGTGTGGCAAGTTTCAACTCTGCTCT 680
689 CTTTGGCATGATGGGACCATCTTCTGCTGGGTTTCTTGGGCGAGCTTCAACTCGCGCC 748
691 GCTGAGAGTCCAATCGAAGAGAAATGCGGTGTTCAACACCTTACTTGTGTAGCAGT 740
749 GACAGCGTGGGGAGTGGGAGCATCGGACCGTGGTCAACATATCTTACTTACTCAGCG 808
741 CAGCGTGGTGACGACCATCTCAGGGTCTATCTTGGCTCACCCCAAGGGAAGATCAGCA 800

Db 809 AAGCACCCTCAGTACTTTTCGCTTGTCAAGCCCTTGTCAAGTGGAGATGGACGACAT 868
QY 801 GACTTATGTGACAGTGGCTGTTGGCAGGAGCGTGGCTGGGTACCTGCTGTCACT 860
Db 869 GGTCCACGCTCCAGAACGACGACACTGGCTGGAGGTGTTGGTGGGACATCAAGTGA 928
QY 861 GATCCCTTCTCGTGGCTTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
Db 929 GATGCTGACACCTTTGGGGCGCTGGCAGCTGGCTTCTGCTGGCTGGAGTGTCTCC 988
QY 921 GGGAGCCAAAGTACCTGCCGGGGGTGTGTAACGAGTGTGGGGATTCGCCACAGTCC 980
Db 989 GGGGTATAGTCTTTTACGCTATCTTGAATCCAGATTAAACTGCAAGACACATGTGG 1048
QY 981 CATGGGCTACACTTTCAGCTTCTGGTCTGCTGGAGAGATCATCTACATTTGCTGCT 1040
Db 1049 TGTTCACACCTTCCATGGGATGCCAGGGTCTCTGGGGGCACTCTGGGAGTCTGT 1108
QY 1041 GGTGCTTGTATACCTGCTGGAGC 1061
Db 1109 TGCACCTGGCCACCCACGAGC 1129
RESULT 13
AAH18728/c
ID AAH18728 standard; cDNA; 3706 BP.
XX
AC AAH18728;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:19002.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 19002; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dr primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 3706 BP; 883 A; 1050 C; 684 G; 1089 T; 0 U; 0 Other;

Query Match 12.2%; Score 152.4; DB 4; Length 3706;
Best Local Similarity 99.4%; Pred. No. 5.4e-33;
Matches 153; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 333 CAGTATTCGGCTGGCCACCATGAGTCTTTGTCGGTGTGATCTCAGTGTGATCTCTCTT 392
Db |||||||
QY 2906 CAGTATTCGGCTGGCCACCATGAGTCTTTGTCGGTGTGATCTCAGTGTGATCTCTT 2847
Db |||||||
QY 393 GGGGAAGGTCAACTTGGCGCAGTGTGGTGTGATGTGCTGGTGGAGTGCACAGCTTTAG 452
Db |||||||
QY 2846 GGGGAAGGTCAACTTGGCGCAGTGTGGTGTGATGTGCTGGTGGAGTGCACAGCTTTAG 2787
Db |||||||
QY 453 CAACTCTAGAGTGTGATCAGTAATATCTTCAAC 486
Db |||||||
QY 2786 CAACTCTAGAGTGTGATCAGTAATATCTTCAAC 2753
Db |||||||

RESULT 14
ABK49216
ID ABK49216 standard; cDNA; 2098 BP.
XX
AC ABK49216;
XX
DT 15-JUL-2002 (first entry)
XX
DE cDNA encoding mouse Rh type C gene (rhcg) protein.
XX
KW RHCG; mouse; non-erythroid Rh type C glycoprotein; chromosome 7; gene;
KW ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 123..1496
FT /tag= a
FT /product= "rhcg protein"
FT polyA_signal 1934..1939
FT /tag= b
XX
FN WO200220719-A2.
XX
PD 14-MAR-2002.
XX
PF 05-SEP-2001; 2001WO-US027503.
XX
PR 07-SEP-2000; 2000US-0230660P.
XX
PA (NYBL-) NEW YORK BLOOD CENT INC.
XX
PI Huang C, Liu Z;
XX
DR WPI; 2002-351774/38.
DR P-PSDB; AAU78998.
XX
PT Nucleic acid sequences encoding novel mammalian nonerythroid Rh type C
PT and glycoproteins which have a characteristic twelve transmembrane domain
PT structure.

XX Claim 3; Fig 1; 53pp; English.

This invention relates to the nucleic acid and protein sequences of novel human and mouse non-erythroid Rh type C glycoprotein (RhCG). The RhCG protein and the mouse homologue (rhcg) have a characteristic 12 transmembrane domain structure and are expressed in kidneys and testis. The invention also comprises a method for antibody that specifically binds an epitope of the glycoprotein and a method for detecting the protein using this antibody. The antibodies of the invention may be used in Western blots, enzyme linked immunosorbent assays (ELISA) or immunohistochemical assays to identify the non-erythroid tissues, particularly kidney and testis, that express the RhCG or Rhcg glycoproteins. The methods are used for detecting an Rhcg or RhCG glycoprotein in a sample. The present sequence represents the cDNA encoding the mouse RhCG Rh type C glycoprotein (RhCG) protein sequence of the invention. The gene encoding this protein is located on murine chromosome 7

SQ Sequence 2098 BP; 458 A; 594 C; 559 G; 487 T; 0 U; 0 Other;

Query Match 12.0%; Score 150.2; DB 6; Length 2098;
Best Local Similarity 47.5%; Pred. No. 1.7e-32;
Matches 446; Conservative 0; Mismatches 493; Indels 0; Gaps 0;
QY 97 CACTATGACGCTTCCTTAGAGGATCAAAAGGGGCTCGTGGCATCTCTCAAGTTGGCCAA 156
Db |||||||
QY 258 CGCAAGAACATCTCCAGCGACGTTGAGAACGAGTTTACTATCGCTACCGAGCTCCAG 317
Db |||||||
QY 157 GATCTGACCGTATGGCGCCATTCGCTTGGCTTCCTCACCCTCGAGTTCCGGAGACAC 216
Db |||||||
QY 318 GATGTACACGCCATGCTTCGTGGGCTTCGCTTCCTCATGACCTTCCTCGAGCGCTAC 377
Db |||||||
QY 217 AGCTGGAGCAGTGTGGCGCTTCAACCTCTTCATGCTGGCGCTTGGTGGAGTGCACAGCTTTAGGCAAC 276
Db |||||||
QY 378 GCTTCAGCGCTGTAGGCTTCACTTCTGCTGGCAGCTTCGGCATCAGTGGGCACTG 437
Db |||||||
QY 277 CTGCTGGACGGCTTCCTGAGCCAGTTCCTCTGGAAGAGTGGTGCATCACACTGTTCAGT 336
Db |||||||
QY 438 CTCATGCGAGGATGGTTCATTACTTTGAAGAAGCCCATTTGCTGAGCGCTCGAGAAC 497
QY 337 ATTTCGGCTGGCCACCATGAGTGTCTTGTGGTGTGATCTCAGTGGATGCTCTCTTGGGG 396
Db |||||||
QY 498 ATCATCCAAAGTGACTTCGTGTGGCATCTTCCTGTGTGGCTTCGGGCGAGCTTCTAGGC 557
Db |||||||
QY 397 AAGGTCAAACCTTGGCGCAGTGTGGTGTGATGTGTGGTGGAGTGCACAGCTTTAGGCAAC 456
Db |||||||
QY 558 AAGGTCAAGCGGATGCAGCTGCTCATATGACCTTCTTCCAGTGAATCTCTCTTACAGTG 617
QY 457 CTGAGGATGGTCAATCAGTAATATCTTCAACACAGACTACCATGAAATGATGACATC 516
Db |||||||
QY 618 AATGAGTTTCATCTCTGAACTGATAGAGGCAAGGATGAGGGGGCTCTATGACCATC 677
QY 517 TAGCTGTTTCGAGCCTATTTTGGGCTGTCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 576
Db |||||||
QY 678 CACACATTTGGCGCCTACTTTGGGCTCACAGTGAACCTGGATCTCTCTACCGAAAAACCTG 737
QY 577 GAGGGAACGAGGATTAAGATCAGACAGCAACCATACCCAGTTTGTCTGCCATGCTGGGC 636
Db |||||||
QY 738 GATCAGAGCAAGCAGACAGAGCTCAGTGTACCACCTCGGACCTTTTCGCCATGATTTGGC 797
QY 637 GCGCTCTTCTTGTGATGTTCTGGCCAAAGTTTCAACTCTGCTGTCTGTCTGTCTGTCTGTCT 696
Db |||||||
QY 798 ACCCTCTTCTTGTGATATCTGGCCCAAGTTTCAATTTCAGCCAGTTCCTTCCACGGAGAT 857
QY 697 GAAAGGAGATGCGGTGTTCAACACCTACTATCTGTAGCAGTCAAGGTTGGTGCAGGCC 756
Db |||||||
QY 858 GCCCAGCACCGAGCGCCCTCAATACCTCTCTCTTGGCAGCGAGTGTCTCAACCA 917
QY 757 ATCTCAGGCTCATCTTGGCTCACCCCAAGGGAGATCAGCAAGACTTATGTGCACAGT 816
Db |||||||
QY 918 GTGACAGTATCCAGTATTTGTACCAAGAGGGCAAGTTGGATATGGTGCACATCCAGAT 977
Db |||||||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2004, 12:40:53 ; Search time 3117 Seconds

(without alignments)
12013.851 Million cell updates/sec

Title: US-09-600-714-41

Perfect score: 1254

Sequence: 1 atgagctctaagtaaccgcg.....atttgctgttgattttaa 1254

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_nam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gsl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	824.4	65.7	994	13	BX406635
2	633.8	50.5	685	13	BX406635
3	629.8	50.2	676	13	BX106517
4	617.2	49.2	685	13	BX406635

C	5	609	48.6	1000	13	BX406634
6	602.2	48.0	48.0	1457	11	AK079335
7	597.4	47.6	47.6	706	13	BU661907
8	560.4	44.7	44.7	631	13	BU662207
9	559	44.6	44.6	632	13	BU665450
10	552	44.0	44.0	613	13	BU664859
11	542.4	43.3	43.3	607	13	BU667533
12	498.6	39.8	39.8	975	11	AK089642
13	427.2	34.1	34.1	612	13	BU663660
14	405.6	32.3	32.3	587	13	BU661598
15	399.2	31.8	31.8	480	13	BU659596
16	398.6	31.8	31.8	683	12	BM719724
17	392.4	31.3	31.3	423	9	AL699045
18	359	28.6	28.6	449	13	BU657883
19	349.2	27.8	27.8	721	13	BY750607
20	343.2	27.4	27.4	551	12	BM254191
21	321.2	25.6	25.6	660	13	BY722497
22	317.6	25.3	25.3	643	12	BM684087
23	314.4	25.1	25.1	701	13	BY734317
24	302.2	24.1	24.1	434	14	NS9044
25	299.6	23.9	23.9	375	13	BU664580
26	294.2	23.5	23.5	499	10	BF603905
27	288.2	23.0	23.0	488	10	BF602079
28	280.4	22.4	22.4	560	29	CG573338
29	273.6	21.8	21.8	515	13	EX529358
30	271.2	21.6	21.6	498	12	BG944752
31	270.4	21.6	21.6	416	14	R10548
32	257.8	20.6	20.6	601	13	BU661774
33	255.8	20.4	20.4	522	29	CG634208
34	253.6	20.2	20.2	343	14	T84327
35	250	19.9	19.9	504	29	CG669296
36	243	19.4	19.4	515	10	BF191605
37	240.8	19.2	19.2	474	13	BY563146
38	240.8	19.2	19.2	476	13	BY559962
39	240.8	19.2	19.2	476	13	BY592375
40	240.2	19.2	19.2	522	29	CG570751
41	239.8	19.1	19.1	501	29	CG558871
42	236.8	18.9	18.9	491	9	AA259801
43	231.8	18.5	18.5	446	10	BB847808
44	230.2	18.4	18.4	461	13	BY557825
45	225.8	18.0	18.0	453	13	BY591481

ALIGNMENTS

RESULT 1	BX406635	994 bp	mRNA	linear	EST 15-MAY-2003
LOCUS	BX406635	Homo sapiens FETAL LIVER	Homo sapiens	CDNA clone	
DEFINITION	CSODM013YA17 5-PRIME, mRNA sequence.				
ACCESSION	BX406635				
VERSION	BX406635.1	GI:30762708			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Li W.B., Gruber C., Jessee J., and Polayes D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 854.r For more information about this cluster, see http://www.genoscope.cns.fr/ cgl-bin/cluster.cgi?seq=CSOAM013AA09QP1&cluster=854.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600				

Paraday Avenue Genoscope sequence ID : CS0AM013AA09P1.

FEATURES

Location/Qualifiers
1. .994
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM013A17"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
ORIGIN
Query Match 65.7%; Score 824.4; DB 13; Length 994;
Best Local Similarity 93.0%; Pred. No. 7.3e-198;
Matches 857; Conservative 16; Mismatches 47; Indels 2; Gaps 1;
QY 1 ATGAGCTCTAAGTACCGCGGTCTGTCGGCGTCTGTCGCCCTCTGGGCGCTTAACACTG 60
Db ATGAGCTCTAAGTACCGCGGTCTGTCGGCGTCTGTCGCCCTCTGGGCGCTTAACACTG 132
QY 61 GAAGCAGCTCTCATCTCCCTCTCTATTTTACCCACTATGACGCTTCTCTTAGAGGAT 120
Db 133 GAAGCAGCTCTCATCTCCCTCTCTATTTTBBBACCACACTATGACGCTTCTCTTAGAGGAT 192
QY 121 CAAAGGGGCTCGGATCTCTATCAAGTTGGCCAGATCTGACCGTGTGGCGGCCATT 180
Db 193 CAARAGGGGCTCGGATCTCTATCAAGTGGCCAGATCTGACCGTGTGGCGGCCATT 252
QY 181 GGCTTGGGCTTCTCCTCAGCTCGAGTTTCGGAGACACAGCTGGAGAGTGGCGCTTCAAC 240
Db 253 GGCTTGGGCTTCTCCTCAGCTCGAGTTTCGGAGACACAGCTGGAGAGTGGCGCTTCAAC 312
QY 241 CTCTTCATGCGGCTGGTGTGTCAGTGGGCAATCTGTCGGAGCGCTTCTCTGAGCCAG 300
Db 313 CTCTTCATGCGGCTGGTGTGTCAGTGGGCAATCTGTCGGAGCGCTTCTCTGAGCCAG 372
QY 301 TTCCCTTCTGGAGGTTGGTCTATCACACTGTTCAGTATTCGGCTGGCCACCATGAGTGT 360
Db 373 TTCCCTTCTGGAGGTTGGTCTATCACACTGTTCAGTATTCGGCTGGCCACCATGAGTGT 432
QY 361 TTGTGGGTCTGATCTAGTGGATGCTCTTTGGGGAAGTCAACTTGGCGCAGTTGGTG 420
Db 433 ATGTGGGTCTGATCTAGTGGATGCTCTTTGGGGAAGTCAACTTGGCGCAGTTGGTG 492
QY 421 GTGATGCTGGTGGAGTGACAGCTTTAGGCACTGAGGATGGTCTATCAGTAATATC 480
Db 493 GTGATGCTGGTGGAGTGACAGCTTTAGGCACTGAGGATGGTCTATCAGTAATATC 552
QY 481 TTCAACACAGACTACCATGAACATGATGACATCTACGTTTTCGACGCTTATTTGGG 540
Db 553 TTCAACACAGACTACCATGAACATGATGACATCTACGTTTTCGACGCTTATTTGGG 612
QY 541 CTGTCTGGGCTGGTGGCTGCGCAAGCTCTACCGAGGGAACGAGGATAAAGATCAG 600
Db 613 CTGTCTGGGCTGGTGGCTGCGCAAGCTCTACCGAGGGAACGAGGATAAAGATCAG 672
QY 601 ACAGCAACGATACCCAGTTGTCTCCATGCTGGGCGCTCTCTTGTGGATGTTCTGG 660
Db 673 AGAGCAACGATACCCAGTTGTCTCCATGCTGGGCGCTCTCTTGTGGATGTTCTGG 732
QY 661 CCAAGTTTCAACTCTGCTCTGAGAACTCCAATCGAAAGGAAGAAAGTCCGCTGTTCAAC 720
Db 733 CCAAGTTTCAACTCTCTGCTGAGAACTCCAATCGAAAGGAAGAAAGTCCGCTGTTCAAC 792
QY 721 ACCTACTGCTGTAGCAGTGGGTTGGTGGAGCAGCATCTCAGGTCATCTTGGCTCAC 780
Db 793 ACCTACTGCTGTAGCAGTGGGTTGGTGGAGCAGCATCTCAGGTCATCTTGGCTCAC 852

QY 781 CCCAAGGAAGATCAGCAAGACTTATGTGCACAGTGCCTGTGTGGAGAGCGGTGGCT 840
Db 853 CCCAAGGAAGATCAGCAAGACTTATGTGCACAGTGCCTGTGTGGAGAGCGGTGGCT 912
QY 841 GTGGGTACCTCG--TGTACCTGATCCCTTCTCCGTGGCTTGCATGCTCGGTCTTG 898
Db 913 GTGGGTACTCGTGTCTACCTGAKCCCTTCTCGGGGGCTTGCASWSGTGCTCGGTCTC 972
QY 899 TGGCTGGCTGATCTCCGTGG 920
Db 973 TTGGCTGGGTGATTTGCAHCGG 994
BU664461 685 bp mRNA linear EST 30-SEP-2002
LOCUS c1117b02.z1 Hembase; Erythroid Precursor Cells (LCB:cl library)
DEFINITION Homo sapiens cDNA clone c1117b02 5', mRNA sequence.
ACCESSION BU664461
VERSION BU664461.1 GI:23376646
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 685)
AUTHORS Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
TITLE Gene Expression in Human Erythroid Precursor Cells
JOURNAL Unpublished (2002)
COMMENT Contact: Jeffery L. Miller
Laboratory of Chemical Biology
National Institute of Diabetes and Digestive and Kidney Diseases
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD 20892, USA
Tel: 301 402 2373
Fax: 301 435 5148
Email: jmf@nih.gov
The 'cl' library was constructed by Alexander Gubin, Ph.D. in the Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or analyses by National Institutes of Health Intramural Sequencing Center (NISC). More information available at:
http://hembase.nidk.nih.gov
Plate: 117 row: b column: 02
Seq primer: 5' lambda-TripleX2 Sequencing Primer.
FEATURES
Location/Qualifiers
1. .685
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="c1117b02"
/sex="unknown"
/tissue_type="blood"
/cell_type="Erythroid Precursor Cells"
/cell_line="Primary Culture of Peripheral Blood Mononuclear Cells"
/dev_stage="Precursor erythroblasts; GPA++"
/lab_host="DH5alpha"
/clone_lib="Hembase; Erythroid Precursor Cells (LCB:cl library)"
/note="Organ: blood; Vector: pTriplex2; Site 1: SfiI; Site 2: SfiI; A complementary DNA (cDNA) library from human erythroid precursor cells was constructed using SMART PCR (polymerase chain reaction) cDNA Library Construction Kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1 umol/L peptide nucleic acid (PNA) oligos
(N-terminal)-biotin-GTC-CAC-CCG-AGG-CTT-G-(C-terminal) and (N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal). Synthesized cDNA was digested with SfiI and size-selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH

QY 541 CTGCTGTGCTGGTGGCTGCGCAAGACCTCTACCGAGGAAACGAGGATGAATGATCAG 600
 Db 568 TTGACTGTGCTGGTGGCTGCGCAAGACCTCTACCGAGGAAACGAGGATGAATGATCAG 627
 QY 601 ACAGCAACGATACCCAGTTGCTGCGCATGCTGGGCGCCCTCTCTTGT 649
 Db 628 AGAGCAACGATACCCAGTTGCTGCGCATGCTGGGCGCCCTCTCTTGT 676

RESULT 4
 BU655978
 LOCUS
 DEFINITION
 sapiens cDNA clone cl04g11 5', mRNA sequence.

ACCESSION
 BU655978
 VERSION
 EST.
 SOURCE
 Homo sapiens (human)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

685 bp mRNA linear EST 30-SEP-2002
 cl04g11.z1 Hembase; Erythroid Precursor Cells (LCB:cl library) Homo
 sapiens cDNA clone cl04g11 5', mRNA sequence.
 BU655978
 GI:23368160
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Gubin, A.N., Lee, Y.T., Bouffard, G.G. and Miller, J.L.
 Gene Expression in Human Erythroid Precursor Cells
 Unpublished (2002)
 Contact: Jeffrey L. Miller
 Laboratory of Chemical Biology
 National Institute of Diabetes and Digestive and Kidney Diseases
 Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
 20892, USA
 Tel: 301 402 2373
 Fax: 301 435 5148
 Email: jm7fenih.gov

The 'cl' library was constructed by Alexander Gubin, Ph.D. in the
 Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
 analyses by National Institutes of Health Intramural Sequencing
 Center (NISC). More information available at:
<http://hembase.nidk.nih.gov>
 Plate: 04 row: 9 column: 11
 Seq primer: 5' lambda-Triplex2 Sequencing Primer.

FEATURES
 source

1..685
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="cl04g11"
 /sex="unknown"
 /tissue_type="blood"
 /cell_type="Erythroid Precursor Cells"
 /cell_line="Primary Culture of Peripheral Blood
 Mononuclear Cells"
 /dev_stage="Precursor erythroblasts; GPA++"
 /lab_host="DH5alpha"
 /clone_lib="Hembase; Erythroid Precursor Cells (LCB:cl
 library)"
 /note="Organ: blood; Vector: pTriplex2; Site 1: SfiI;
 Site 2: SfiI; A complementary DNA (cDNA) library from
 human erythroid precursor cells was constructed using
 SMART PCR (polymerase chain reaction) cDNA Library
 Construction Kit (Clontech, Palo Alto, CA) according to
 the manufacturer's directions, but with slight
 modifications. Briefly, reverse transcription was
 performed in the presence of 1 umol/L peptide nucleic acid
 (PNA) oligos
 (N-terminal)-biotin-CTC-CAC-COG-AGG-CTT-G-(C-terminal) and
 (N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal)
 . Synthesized cDNA was digested with SfiI and
 size-selected on a 1% agarose gel (>800bp). Large-scale
 sequencing of the library was performed by the NIH
 Intramural Sequencing Center (NISC;
<http://www.nisc.nih.gov/>).

ORIGIN

Query Match 49.2%; Score 617.2; DB 13; Length 685;
 Best Local Similarity 98.0%; Pred. No. 2.3e-145;
 Matches 625; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 QY 1 ATGAGCTCTAAGTACCGCGGTCTGTCTCGGCGCTGCTGCCCTCTGGGCGCCCTAACTG 60
 Db 48 ATGAGCTCTAAGTACCGCGGTCTGTCTCGGCGCTGCTGCCCTCTGGGCGCCCTAACTG 107
 QY 61 GAAGCAGCTCTCATTTCTCTCTCTTTTACCCACTATGAGCGCTTCTTTAGAGGAT 120
 Db 108 GAAGCAGCTCTCATTTCTCTCTCTTTTACCCACTATGAGCGCTTCTTTAGAGGAT 167
 QY 121 CAAAGGGGCTGTGTCATCTCAAGTTGGCCAAAGATCTGACCGCTGATGGCGGCATT 180
 Db 168 CAAAGGGGCTGTGTCATCTCAAGTTGGCCAAAGATCTGACCGCTGATGGCGGCATT 227
 QY 181 GGCTTGGGCTTCTACCTCGAGTTTCCGAGACACAGCTGGAGCATGTGGCTTCAAC 240
 Db 228 GGCTTGGGCTTCTACCTCGAGTTTCCGAGACACAGCTGGAGCATGTGGCTTCAAC 287
 QY 241 CTCTTCATGCTGCGCTTGGTGTGAGTGGCAATCTCTGTCGACCGCTTCTCTGAGCCAG 300
 Db 288 CTCTTCATGCTGCGCTTGGTGTGAGTGGCAATCTCTGTCGACCGCTTCTCTGAGCCAG 347
 QY 301 TTCCCTTCTGGGAAGGTGGTCAACACTTTCAGTATTGGTGGCCACCATGAGTGTCT 360
 Db 348 TTCCCTTCTGGGAAGGTGGTCAACACTTTCAGTATTGGTGGCCACCATGAGTGTCT 407
 QY 361 TTGTCGGTGTGATCTCAGTGGATGTCTTGGGGAAGTCAACTTGGCGCAGTTGGTG 420
 Db 408 GTGTCGGTGTGATCTCAGCGGTGTCTTCTGGGGAAGTCAACTTGGCGCAGTTGGTG 467
 QY 421 GTCATGCTGCTGGGAGGTGACAGCTTTAGGCAACCTGAGGATGGTCAATGATATC 480
 Db 468 GTCATGCTGCTGGGAGGTGACAGCTTTAGGCAACCTGAGGATGGTCAATGATATC 527
 QY 481 TTCAACACAGCTACCAACATGAACATGATGACATCTACCTGTTCGAGCCTATTTGGG 540
 Db 528 TTCAACACAGCTACCAACATGAACATGATGACATCTACCTGTTCGAGCCTATTTGGG 587
 QY 541 CTGCTGTGCTGGTGGCTGCCAAGCCTCTACCGAGGAAACGAGGATGAATGATCAG 600
 Db 588 CTGACTGTGCGCTGGTGGCTGCCAAGCCTCTACCGAGGAAACGAGGATGAATGATCAG 647
 QY 601 ACAGCAACGATACCCAGTTGCTGCGCATGCTGGGCGC 638
 Db 648 AGAGCAACGATACCCAGTTGCTGCGCATGCTGGGCGC 685

RESULT 5

BU655978

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Comment

1000 bp mRNA linear EST 13-MAY-2003
 BU655978 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
 CS0DM013YA17 3-PRIME, mRNA sequence.

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Comment

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: sequef@genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 854.r

For

more information about this cluster, see

[illegible]

RESULT 9	ACCESSION	ORGANISM
BU665450	VERSION	
LOCUS	KEYWORDS	
DEFINITION	SOURCE	

BU665450.1 632 bp mRNA linear EST 30-SEP-2002
 cl131d04.z1 Hembase; Erythroid Precursor Cells (UCB:cl library)
 Homo sapiens cDNA clone cl131d04 5', mRNA sequence.
 BU665450
 BU665450.1 GI:23377637
 EST.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 632)
 Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
 Gene Expression in Human Erythroid Precursor Cells
 Unpublished (2002)
 Contact: Jeffery L. Miller
 Laboratory of Chemical Biology
 National Institute of Diabetes and Digestive and Kidney Diseases
 Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
 20892, USA

Tel: 301 402 2373
Fax: 301 435 5148
Email: jm7@nih.gov
The 'cl' library was constructed by Alexander Gubin, Ph.D. in the Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or analyses by National Institutes of Health Intramural Sequencing Center (NISC). More information available at:
<http://hembase.nidk.nih.gov>
Plate: 131 row: d column: 04

FEATURES	SOURCE
1. <i>Age</i>	1. <i>Age</i>
2. <i>Gender</i>	2. <i>Gender</i>
3. <i>Marital status</i>	3. <i>Marital status</i>
4. <i>Education</i>	4. <i>Education</i>
5. <i>Income</i>	5. <i>Income</i>
6. <i>Occupation</i>	6. <i>Occupation</i>
7. <i>Religion</i>	7. <i>Religion</i>
8. <i>Political affiliation</i>	8. <i>Political affiliation</i>
9. <i>Health status</i>	9. <i>Health status</i>
10. <i>Travel history</i>	10. <i>Travel history</i>
11. <i>Family size</i>	11. <i>Family size</i>
12. <i>Home ownership</i>	12. <i>Home ownership</i>
13. <i>Employment status</i>	13. <i>Employment status</i>
14. <i>Language spoken at home</i>	14. <i>Language spoken at home</i>
15. <i>Number of children</i>	15. <i>Number of children</i>
16. <i>Years in current residence</i>	16. <i>Years in current residence</i>
17. <i>Number of previous residences</i>	17. <i>Number of previous residences</i>
18. <i>Years since last move</i>	18. <i>Years since last move</i>
19. <i>Number of pets</i>	19. <i>Number of pets</i>
20. <i>Years since last pet acquisition</i>	20. <i>Years since last pet acquisition</i>
21. <i>Number of vehicles</i>	21. <i>Number of vehicles</i>
22. <i>Years since last vehicle purchase</i>	22. <i>Years since last vehicle purchase</i>
23. <i>Number of trips abroad</i>	23. <i>Number of trips abroad</i>
24. <i>Years since last trip abroad</i>	24. <i>Years since last trip abroad</i>
25. <i>Number of family members</i>	25. <i>Number of family members</i>
26. <i>Years since last family member addition</i>	26. <i>Years since last family member addition</i>
27. <i>Number of siblings</i>	27. <i>Number of siblings</i>
28. <i>Years since last sibling birth</i>	28. <i>Years since last sibling birth</i>
29. <i>Number of children under 18</i>	29. <i>Number of children under 18</i>
30. <i>Years since last child birth</i>	30. <i>Years since last child birth</i>
31. <i>Number of children 18 and over</i>	31. <i>Number of children 18 and over</i>
32. <i>Years since last child 18 and over</i>	32. <i>Years since last child 18 and over</i>
33. <i>Number of children in college</i>	33. <i>Number of children in college</i>
34. <i>Years since last child in college</i>	34. <i>Years since last child in college</i>
35. <i>Number of children in military</i>	35. <i>Number of children in military</i>
36. <i>Years since last child in military</i>	36. <i>Years since last child in military</i>
37. <i>Number of children in workforce</i>	37. <i>Number of children in workforce</i>
38. <i>Years since last child in workforce</i>	38. <i>Years since last child in workforce</i>
39. <i>Number of children in home</i>	39. <i>Number of children in home</i>
40. <i>Years since last child in home</i>	40. <i>Years since last child in home</i>
41. <i>Number of children in care</i>	41. <i>Number of children in care</i>
42. <i>Years since last child in care</i>	42. <i>Years since last child in care</i>
43. <i>Number of children in foster care</i>	43. <i>Number of children in foster care</i>
44. <i>Years since last child in foster care</i>	44. <i>Years since last child in foster care</i>
45. <i>Number of children in adoption</i>	45. <i>Number of children in adoption</i>
46. <i>Years since last child in adoption</i>	46. <i>Years since last child in adoption</i>
47. <i>Number of children in custody</i>	47. <i>Number of children in custody</i>
48. <i>Years since last child in custody</i>	48. <i>Years since last child in custody</i>
49. <i>Number of children in juvenile detention</i>	49. <i>Number of children in juvenile detention</i>
50. <i>Years since last child in juvenile detention</i>	50. <i>Years since last child in juvenile detention</i>
51. <i>Number of children in prison</i>	51. <i>Number of children in prison</i>
52. <i>Years since last child in prison</i>	52. <i>Years since last child in prison</i>
53. <i>Number of children in mental health care</i>	53. <i>Number of children in mental health care</i>
54. <i>Years since last child in mental health care</i>	54. <i>Years since last child in mental health care</i>
55. <i>Number of children in substance abuse treatment</i>	55. <i>Number of children in substance abuse treatment</i>
56. <i>Years since last child in substance abuse treatment</i>	56. <i>Years since last child in substance abuse treatment</i>
57. <i>Number of children in residential care</i>	57. <i>Number of children in residential care</i>
58. <i>Years since last child in residential care</i>	58. <i>Years since last child in residential care</i>
59. <i>Number of children in community care</i>	59. <i>Number of children in community care</i>
60. <i>Years since last child in community care</i>	60. <i>Years since last child in community care</i>
61. <i>Number of children in foster family</i>	61. <i>Number of children in foster family</i>
62. <i>Years since last child in foster family</i>	62. <i>Years since last child in foster family</i>
63. <i>Number of children in adoption family</i>	63. <i>Number of children in adoption family</i>
64. <i>Years since last child in adoption family</i>	64. <i>Years since last child in adoption family</i>
65. <i>Number of children in custody family</i>	65. <i>Number of children in custody family</i>
66. <i>Years since last child in custody family</i>	66. <i>Years since last child in custody family</i>
67. <i>Number of children in juvenile detention family</i>	67. <i>Number of children in juvenile detention family</i>
68. <i>Years since last child in juvenile detention family</i>	68. <i>Years since last child in juvenile detention family</i>
69. <i>Number of children in prison family</i>	69. <i>Number of children in prison family</i>
70. <i>Years since last child in prison family</i>	70. <i>Years since last child in prison family</i>
71. <i>Number of children in mental health care family</i>	71. <i>Number of children in mental health care family</i>
72. <i>Years since last child in mental health care family</i>	72. <i>Years since last child in mental health care family</i>
73. <i>Number of children in substance abuse treatment family</i>	73. <i>Number of children in substance abuse treatment family</i>
74. <i>Years since last child in substance abuse treatment family</i>	74. <i>Years since last child in substance abuse treatment family</i>
75. <i>Number of children in residential care family</i>	75. <i>Number of children in residential care family</i>
76. <i>Years since last child in residential care family</i>	76. <i>Years since last child in residential care family</i>
77. <i>Number of children in community care family</i>	77. <i>Number of children in community care family</i>
78. <i>Years since last child in community care family</i>	78. <i>Years since last child in community care family</i>
79. <i>Number of children in foster family care</i>	79. <i>Number of children in foster family care</i>
80. <i>Years since last child in foster family care</i>	80. <i>Years since last child in foster family care</i>
81. <i>Number of children in adoption family care</i>	81. <i>Number of children in adoption family care</i>
82. <i>Years since last child in adoption family care</i>	82. <i>Years since last child in adoption family care</i>
83. <i>Number of children in custody family care</i>	83. <i>Number of children in custody family care</i>
84. <i>Years since last child in custody family care</i>	84. <i>Years since last child in custody family care</i>
85. <i>Number of children in juvenile detention family care</i>	85. <i>Number of children in juvenile detention family care</i>
86. <i>Years since last child in juvenile detention family care</i>	86. <i>Years since last child in juvenile detention family care</i>
87. <i>Number of children in prison family care</i>	87. <i>Number of children in prison family care</i>
88. <i>Years since last child in prison family care</i>	88. <i>Years since last child in prison family care</i>
89.	

Qy	1	ATGAGCTCTAAGTACCGCGGTCTGTGTCGGCGCTGCTGCCCCCTCTGGCGCCCTTAACA	CTG 60
Db	69	ATGAGCTCTAAGTACCGCGGTCTGTGTCGGCGCTGCTGCCCCCTCTGGCGCCCTTAACA	CTG 128
Qy	61	GAAGCAGCTCTCATTTCTCTCTCTATTTTTTTTACCCACTATGACGCTTCTCTAGAGGAT	120
Db	129	GAAGCAGCTCTCATTTCTCTCTCTATTTTTTTTACCCACTATGACGCTTCTCTAGAGGAT	188
Qy	121	CAAAAGGGGCTCGTGGCATCCTATCAAGTTGGCCAAAGATCTGACCGTGATGGCGCCANT	180
Db	189	CAAAAGGGGCTCGTGGCATCCTATCAAGTTGGCCAAAGATCTGACCGTGATGGCGCCANT	248
Qy	181	GGCTTGGGCTTCTCACTCGAGTTTCCGGAGACACAGCTGGAGGAGTGTGGCCCTTCAAC	240
Db	249	GGCTTGGGCTTCTCACTCGAGTTTCCGGAGACACAGCTGGAGGAGTGTGGCCCTTCAAC	308
Qy	241	CTCTTCATGCTGGCGCTTGGTGTGTCAGTGGGCAATCCTGCTGGAGCGCTTCTCGAGCAG	300
Db	309	CTCTTCATGCTGGCGCTTGGTGTGTCAGTGGGCAATCCTGCTGGAGCGCTTCTCGAGCAG	368
Qy	301	TTCCCTTCTGGGAAGGTGGTCAATCACACTGTTCAGTATTCGGCTGGCCACCACATGAGTCT	360
Db	369	TTCCCTTCTGGGAAGGTAGTCAATCACACTGTTCAGTATTCGGCTGGCCACCACATGAGTCT	428
Qy	361	TTGTCGGTGTGATCTCAGTGGATGTGCTTCTGGGGAAGGTCAACTGGCGCAGTTGTGTG	420
Db	429	TTGTCGGTGTGATCTCAGTGGATGTGCTTCTGGGGAAGGTCAACTGGCGCANTTGTGTG	488
Qy	421	GTGATGTGCTGGTGGAGGTGCACAGCTTTAGGCAACCTGAGATGGTTCATCAGTAATATC	480
Db	489	GTGATGTGCTGGTGGAGGTGCACAGCTTTAGGCAACCTGAGATGGTTCATCAGTAATATC	548
Qy	481	TTCAACACAGACTACCAATGAAATGATGACATCTACGTGTTGGCAGCCTATTTTGGG	540
Db	549	TTCAACACAGACTACCAATGAAATGATGACATCTACGTGTTGGCAGCCTATTTTGGG	608
Qy	541	CTGTCTGTGGCCTGGTGCTGCC	563
Db	609	CTGTCTGTGGCCTGGTGCTGCC	631

ORIGIN

Query Match 44.6%; Score 559; DB 13; Length 632;
Best Local Similarity 98.3%; Pred. No. 1.3e-130;


```

Db 118 GAAGAGCTCTCATTTCTCTCTCTATTTTACCACATAGACGCTTCCTTAGAGAT 177
QY 121 CAAAAGGGCTCGTGCACTCTATCAAGTTGGCCAAAGATCTACCGTGTATGCGGCCATT 180
Db 178 CAAAAGGGCTCGTGCACTCTATCAAGTTGGCCAAAGATCTACCGTGTATGCGGCCATT 237
QY 181 GCGTTGGGCTTCTCACTCTGAGTTTCCGGAGACACAGCTGGAGCAGTGGCCCTTCAAC 240
Db 238 GCGTTGGGCTTCTCACTCTGAGTTTCCGGAGACACAGCTGGAGCAGTGGCCCTTCAAC 297
QY 241 CTCCTTCATCTGGCGCTTGGTGAGTGGCAATCTCTGACGCGCTTCCTGAGCCAG 300
Db 298 CTCCTTCATCTGGCGCTTGGTGAGTGGCAATCTCTGACGCGCTTCCTGAGCCAG 357
QY 301 TTCCTTCTGGGAAGTGTGTATCACTGTGTCAAGTATTCGGTGGCCACCATGAGTGT 360
Db 358 TTCCTTCTGGGAAGTGTGTATCACTGTGTCAAGTATTCGGTGGCCACCATGAGTGT 417
QY 361 TTGTGGTGTCTGATCTCACTGATGCTGTCTTGGGGAAGTCAACTTGGCGGAGTTGGT 420
Db 418 AGCTGGTGTCTGATCTCACTGATGCTGTCTTGGGGAAGTCAACTTGGCGGAGTTGGT 477
QY 421 GTGATGGTCTGCTGGTGAGTGACAGCTTTAGGCAACCTGAGGATGGTCACTAGTAATATC 480
Db 478 GTGATGGTCTGCTGGTGAGTGACAGCTTTAGGCAACCTGAGGATGGTCACTAGTAATATC 537
QY 481 TTCACACAGACTACACATGACATGATGACATCTAGTGTTCGACGCTATTTTGGG 540
Db 538 TTCACACAGACTACACATGACATGATGACATCTAGTGTTCGACGCTATTTTGGG 597
QY 541 CTGTCTGTGGCTGTGGCTGCTGCAAGGCTCTACC 575
Db 598 CTGACTGTGGCTGTGGCTGCTGCAAGGCTCTACC 632

```

```

RESULT 10
BU664659
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BU664659 613 bp mRNA linear EST 30-SEP-2002
c1120e10.z1 Hembase; Erythroid Precursor Cells (LCB:cl library)
Homo sapiens cDNA clone c1120e10 5', mRNA sequence.
ACCESSION
BU664659
EST.
BU664659.1 GI:23376845
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
Gene Expression in Human Erythroid Precursor Cells
Unpublished (2002)
Contact: Jeffery L. Miller
Laboratory of Chemical Biology
National Institute of Diabetes and Digestive and Kidney Diseases
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
20892, USA
Tel: 301 402 2373
Fax: 301 435 5148
Email: jmf7@nih.gov
The 'cl' library was constructed by Alexander Gubin, Ph.D. in the
Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
analyses by National Institutes of Health Intramural Sequencing
Center (NISC). More information available at:
http://hembase.nidddk.nih.gov
Plate: 120 row: e column: 10
Seq primer: 5' lambda-Triplex2 Sequencing Primer.
Location/Qualifiers
1..613
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="c1120e10"

```

```

FEATURES
Source

```

```

/sex="unknown"
/tissue_type="blood"
/cell_type="Erythroid Precursor Cells"
/cell_line="Primary Culture of Peripheral Blood Mononuclear Cells"
/dev_stage="Precursor erythroblasts; GPA++"
/lab_host="DH5alpha"
/clone_lib="Hembase; Erythroid Precursor Cells (LCB:cl library)"
/Note="Organ: blood; Vector: pTriplex2; Site: 1: SfiI; Site 2: SfiI; A complementary DNA (cDNA) library from human erythroid precursor cells was constructed using SMART PCR (polymerase chain reaction) cDNA library Construction kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1 umol/L peptide nucleic acid (PNA) oligos (N-terminal)-biotin-GTC-CAC-CCG-AGG-CTT-G-(C-terminal) and (N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal). Synthesized cDNA was digested with SfiI and size-selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH Intramural Sequencing Center (NISC; Http://www.nisc.nih.gov/)."

```

ORIGIN

```

Query Match 44.0%; Score 552; DB 13; Length 613;
Best Local Similarity 99.6%; Pred. No. 7.7e-129;
Matches 552; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTCTTAGTACCGCGGCTCTGTCGGCGCTGCTGCCCTCTGGGCCCTAACACTG 60
Db 60 ATGAGCTCTTAGTACCGCGGCTCTGTCGGCGCTGCTGCCCTCTGGGCCCTAACACTG 119
QY 61 GAAGCAGCTCTCATTTCTCTCTATTTTACCACATGATGAGCTTCCTTAGAGAT 120
Db 120 GAAGCAGCTCTCATTTCTCTCTATTTTACCACATGATGAGCTTCCTTAGAGAT 179
QY 121 CAAAAGGGCTCGTGGCATCTCATCAAGTTGGCCAAAGATCTGACCGTGTATGCGGCCATT 180
Db 180 CAAAAGGGCTCGTGGCATCTCATCAAGTTGGCCAAAGATCTGACCGTGTATGCGGCCATT 239
QY 181 GCGTTGGGCTTCTCACTCTGAGTTTCCGGAGACACAGCTGGAGCAGTGGCCCTTCAAC 240
Db 240 GCGTTGGGCTTCTCACTCTGAGTTTCCGGAGACACAGCTGGAGCAGTGGCCCTTCAAC 299
QY 241 CTCCTTCATGCTGGCGCTTGGTGTGAGTGGGCAATCTCTGTCGACGGCTTCCTGAGCCAG 300
Db 300 CTCCTTCATGCTGGCGCTTGGTGTGAGTGGGCAATCTCTGTCGACGGCTTCCTGAGCCAG 359
QY 301 TTCCTTCTGGGAAGTGTGTATCACTGTTCAGTATTCGGCTGGCCACCATGAGTGT 360
Db 360 TTCCTTCTGGGAAGTGTGTATCACTGTTCAGTATTCGGCTGGCCACCATGAGTGT 419
QY 361 TTGTGGTGTCTGATCTCACTGATGATCTCTTGGGGAAGTCAACTTGGCGCAGTTGGT 420
Db 420 TTGTGGTGTCTGATCTCACTGATGATCTCTTGGGGAAGTCAACTTGGCGCAGTTGGT 479
QY 421 GTGATGGTGTGGTGGAGGTGACAGCTTTAGGCAACCTGAGGATGGTCACTAGTAATATC 480
Db 480 GTGATGGTGTGGTGGAGGTGACAGCTTTAGGCAACCTGAGGATGGTCACTAGTAATATC 539
QY 481 TTCACACAGACTACACATGACATGATGACATCTAGTGTTCGCGGCCCTATTTTGGG 540
Db 540 TTCACACAGACTACACATGACATGATGACATCTAGTGTTCGCGGCCCTATTTTGGG 599
QY 541 CTGTCTGTGGCTGTG 554
Db 600 CTGTCTGTGGCTGTG 613

```

RESULT 11

```

BU657533
LOCUS          607 bp      mRNA      linear      EST 30-SEP-2002
DEFINITION    BU657533.1 Hembase; Erythroid Precursor Cells (LCB:cl library) Homo
               sapiens cDNA clone cl25d06 5', mRNA sequence.
ACCESSION     BU657533
VERSION       BU657533.1 GI:23369715
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 607)
AUTHORS       Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
TITLE         Gene Expression in Human Erythroid Precursor Cells
JOURNAL       Unpublished (2002)
COMMENT       Contact: Jeffery L. Miller
               Laboratory of Chemical Biology
               National Institute of Diabetes and Digestive and Kidney Diseases
               Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
               20892, USA
               Tel: 301 402 2373
               Fax: 301 435 5148
               Email: jm7f@nih.gov
               The 'cl' library was constructed by Alexander Gubin, Ph.D. in the
               Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
               analyses by National Institutes of Health Intramural Sequencing
               Center (NISC). More information available at:
               http://hembase.nidk.nih.gov
               Plate: 25 row; d column: 06
               Seq primer: 5' lambda-Triplex2 Sequencing Primer.
FEATURES             Location/Qualifiers
     source           1..607
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="cl25d06"
                     /sex="unknown"
                     /tissue_type="blood"
                     /cell_type="Erythroid Precursor Cells"
                     /cell_line="Primary Culture of Peripheral Blood
                     Mononuclear Cells"
                     /dev_stage="Precursor erythroblasts; GPA++"
                     /lab_host="DH5alpha"
                     /clone_lib="Hembase; Erythroid Precursor Cells (LCB:cl
                     library)"
                     /note="Organ: blood; Vector: pTriplex2; Site_1: Sfil;
                     Site_2: Sfil; A complementary DNA (cDNA) library from
                     human erythroid precursor cells was constructed using
                     SMART PCR (polymerase chain reaction) cDNA Library
                     Construction Kit (Clontech, Palo Alto, CA) according to
                     the manufacturer's directions, but with slight
                     modifications. Briefly, reverse transcription was
                     performed in the presence of 1 umol/L peptide nucleic acid
                     (PNA) oligos
                     (N-terminal)-biotin-GTC-CAC-CAG-CTT-G- (C-terminal) and
                     (N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A- (C-terminal)
                     . Synthesized cDNA was digested with Sfil and
                     size-selected on a 1% agarose gel (>800bp). Large-scale
                     sequencing of the library was performed by the NIH
                     Intramural Sequencing Center (NISC;
                     Http://www.nisc.nih.gov/)."
ORIGIN
Query Match          43.3%; Score 542.4; DB 13; Length 607;
Best Local Similarity 98.0%; Pred. No. 2.1e-126;
Matches 549; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 ATGAGCTCTAAGTACCGCGTCTGTCCGCGGTGCTGCGCCCTCTGCGGCCCTAAACACTG 60
Db 48 ATGAGCTCTAAGTACCGGTGTGTGTGTCGCGGTGCTGCGCCCTCTGCGGCCCTAAACACTG 107
Qy 61 GAGCAGCTTCATCTCCCTCTCTATTTTTTACCACATGATGACGGTTCCTTAGAGGAT 120

```

```

Db 108 GAGCAGCTCTCATCTCTCTCTCTCTATTTTTTACCACATGATGACGGCTTCCTTAGAGGAT 167
Qy 121 CAAAAGGGGCTCTGGCATCTCTATCAAGTTGGGCAAGATCTGACCGTGATGGGGCCATT 180
Db 168 CAAAAGGGGCTCTGGCATCTCTATCAAGTTGGGCAAGATCTGACCGTGATGGGGCCATT 227
Qy 181 GGCTTTGGGCTTCTCCTCACCTCGAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCAAC 240
Db 228 GGGTTGGGCTTCTCCTCACCTCGAGTTTCCGGAGACACAGCTGGAGCAGTGTGGGCTTCAC 287
Qy 241 CTCCTTCATGCTGGCGCTTGGTGTGCACTGGGCAATCTCTGCTGGACGGCTTCTTCGAGCCAG 300
Db 288 CTCCTTCATGCTGGCGCTTGGTGTGCACTGGGCAATCTCTGCTGGACGGCTTCTTCGAGCCAG 347
Qy 301 TTCCCTTCTGGGAAGTGTGTCATCACACTGTTTCACTATTCGGCTGGCCACCATGAGTGTCT 360
Db 348 TTCCCTTCTGGGAAGTGTGTCATCACACTGTTTCACTATTCGGCTGGCCACCATGAGTGTCT 407
Qy 361 TTGCTCGGTGCTGATCTCAGTGGATGCTGCTCTGGGGAAGGTCAACTTGGGCGAGTTGGTG 420
Db 408 ATGTCGGTGTGATCTCAGCGGGTGTGCTCTGGGGAAGGTCAACTTGGGCGAGTTGGTG 467
Qy 421 GTCATGGTGTGCTGGAGGTGACAGCTTTAGGCAACCTGAGGATGTCATCAGTAATATC 480
Db 468 GTGATGGTGTGCTGGAGGTGACAGCTTTAGGCAACCTGAGGATGTCATCAGTAATATC 527
Qy 481 TTCAACACAGACTACCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 528 TTCAACACAGACTACCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 587
Qy 541 CTGCTGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 560
Db 588 CTGCTGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 607

RESULT 12
AK089642
LOCUS          975 bp      mRNA      linear      HTC 20-SEP-2003
DEFINITION    Mus musculus activated spleen cDNA, RIKEN full-length enriched
               library, clone:F83006J05 Product:Rhesus blood group CE and D, full
               insert sequence.
ACCESSION     AK089642
VERSION       AK089642.1 GI:26354620
KEYWORDS      HTC; CAP trapper.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1
AUTHORS       Carninci,P. and Hayashizaki,Y.
TITLE         High-efficiency full-length cDNA cloning
JOURNAL       Meth. Enzymol. 303, 19-44 (1999)
MEDLINE       99279253
PUBMED        10349636
REFERENCE     2
AUTHORS       Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
               Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE         Normalization and subtraction of cap-trapper-selected cDNAs to
               prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL       Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE       20499374
PUBMED        11042159
REFERENCE     3
AUTHORS       Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
               Konno,H., Akiyama,J., Nishi,K., Kikunai,T., Tashiro,H., Itoh,M.,
               Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
               Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
               Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
               Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
               Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE         RIKEN integrated sequence analysis (RISA) system--384-Format
               sequencing pipeline with 384 multicapillary sequencer
JOURNAL       Genome Res. 10 (11), 1757-1771 (2000)

```



```

RESULT 13
BU663660          612 bp      mRNA      linear      EST 30-SEP-2002
LOCUS             cl104e12.z1 Hembase; Erythroid Precursor Cells (LCB:cl library)
DEFINITION        Homo sapiens cDNA clone cl104e12 5', mRNA sequence.
ACCESSION         BU663660
VERSION           BU663660.1 GI:23375845
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE            Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
JOURNAL          Gene Expression in Human Erythroid Precursor Cells
COMMENT          Unpublished (2002)
Contact: Jeffery L. Miller
Laboratory of Chemical Biology
National Institute of Diabetes and Digestive and Kidney Diseases
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
20892, USA
Tel: 301 402 2373
Fax: 301 435 5148
Email: jm7@nih.gov
The 'cl' library was constructed by Alexander Gubin, Ph.D. in the
Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
analyses by National Institutes of Health Intramural Sequencing
Center (NISC). More information available at:
http://hembase.nidk.nih.gov
Plate: 104 row: e column: 12
Seq primer: 5' lambda-Triplex2 Sequencing Primer.
FEATURES             Location/Qualifiers
     1..612
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="cl104e12"
     /sex="unknown"
     /tissue_type="blood"
     /cell_type="Erythroid Precursor Cells"
     /cell_line="Primary Culture of Peripheral Blood
Mononuclear Cells"
     /dev_stage="Precursor erythroblasts; GPA++"
     /lab_host="DH5alpha"
     /clone_lib="Hembase; Erythroid Precursor Cells (LCB:cl
library)"
     /notes="Organ: blood; Vector: pTriplex2; Site_1: Sfil;
Site_2: Sfil; A complementary DNA (cDNA) library from
human erythroid precursor cells was constructed using
SMART PCR (polymerase chain reaction) cDNA Library
Construction Kit (Clontech, Palo Alto, CA) according to
the manufacturer's directions, but with slight
modifications. Briefly, reverse transcription was
performed in the presence of 1 umol/L peptide nucleic acid
(PNA) oligos
(N-terminal)-biotin-GTC-CAC-CGG-AAG-CTT-G-(C-terminal) and
(N-terminal)-biotin-C(T/C)T-GAA-CTT-CTC-AGG-A-(C-terminal)
. Synthesized cDNA was digested with Sfil and
size-selected on a 1% agarose gel (>800bp). Large-scale
sequencing of the library was performed by the NIH
Intramural Sequencing Center (NISC;
http://www.nisc.nih.gov/)."
ORIGIN
Query Match          34.1%; Score 427.2; DB 13; Length 612;
Best Local Similarity 89.6%; Pred. No. 3.6e-97;
Matches 489; Conservative 0; Mismatches 13; Indels 44; Gaps 1;
QY      6  CTCCTAAGTACCGCGGTCTGTCGGCGCTGCTGCCCTCTGGGCCCTACACTGGAAGC 65
Db      67  CTCCTAAGTACCGCGGTCTGTCGGCGCTGCTGCCCTCTGGGCCCTACACTGGAAGC 126
QY      66  AGCTCTCANTCTCTCTTCTATTATTTTACCCACTATGAGCTTCCTTAGAGGATCAAAA 125

```

```

Db      127  AGCTCTCANTCTCTCTTCTATTATTTTACCCACTATGAGCTTCCTTAGAGGATCAAAA 186
QY      126  GGGGCTCGTGGGATCTCTATC----- 145
Db      187  GGGGCTCGTGGCATCTCTATCAAGGGGCTGGAAAGTCTGCATGCTGTCTATAAAATCCAGAA 246
QY      146  ----AAAGTTGGCCCAAGATCTGACCGTGGGGCCATTGGCTGGGCTTCCTCACCTCG 201
Db      247  CCAGAAAGTGGCCCAAGATCTGACCGTGGGGCCCTTGGCTGGGCTTCCTCACCTCA 306
QY      202  AGTTTCGGGAGACACAGCTGGAGCAGTGTGGCTTCAACCTCTTTCATGCTGGCGCTGGT 261
Db      307  AATTTCGGGAGACACAGCTGGAGCAGTGTGGCTTCAACCTCTTTCATGCTGGCGCTGGT 366
QY      262  GTGCAGTGGGCAATCTCTGTGGACGGCTTCTGAGCCAGTTCCTCTTGGGAAAGTGGTC 321
Db      367  GTGCAGTGGGCAATCTCTGTGGACGGCTTCTGAGCCAGTTCCTCTTGGGAAAGTGGTC 426
QY      322  ATCACACTGTTCAGTATTTCGGCTGGCCACCATGAGTCTTGTGCGTGTGATCTCAGTG 381
Db      427  ATCACACTGTTCAGTATTTCGGCTGGCCACCATGAGTCTTGTGCGTGTGATCTCAGTG 486
QY      382  GATGCTGTCTTGGGGAAGGTCAACTTGGCGCAGTTCGTGGTGTGATGCTGTGGAGGTG 441
Db      487  GGTGCTGTCTTGGGGAAGGTCAACTTGGCGCAGTTCGTGGTGTGATGCTGTGGAGGTG 546
QY      442  ACAGCTTTAGGCACCTGAGGATGTCATCAGTAATATCTTCAACACAGACTACACATG 501
Db      547  ACAGCTTTATGCACCTGAGGATGTCATCAGTAATATCTTCAACACAGACTACACATG 606
QY      502  AACATG 507
Db      607  AACCTG 612

RESULT 14
BU661598          587 bp      mRNA      linear      EST 30-SEP-2002
LOCUS             cl74c01.z1 Hembase; Erythroid Precursor Cells (LCB:cl library) Homo
DEFINITION        sapiens cDNA clone cl74c01 5', mRNA sequence.
ACCESSION         BU661598
VERSION           BU661598.1 GI:23373780
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE            Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
JOURNAL          Gene Expression in Human Erythroid Precursor Cells
COMMENT          Unpublished (2002)
Contact: Jeffery L. Miller
Laboratory of Chemical Biology
National Institute of Diabetes and Digestive and Kidney Diseases
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
20892, USA
Tel: 301 402 2373
Fax: 301 435 5148
Email: jm7@nih.gov
The 'cl' library was constructed by Alexander Gubin, Ph.D. in the
Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
analyses by National Institutes of Health Intramural Sequencing
Center (NISC). More information available at:
http://hembase.nidk.nih.gov
Plate: 74 row: c column: 01
Seq primer: 5' lambda-Triplex2 Sequencing Primer.
FEATURES             Location/Qualifiers
     1..587
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="cl74c01"

```

```

/sex="unknown"
/tissue_type="blood"
/cell_type="Erythroid Precursor Cells"
/cell_line="Primary Culture of Peripheral Blood Mononuclear Cells"
/dev stage="Precursor erythroblasts; GPA++"
/lab_host="DH5alpha"
/clone_lib="Hembase; Erythroid Precursor Cells (LCB:cl library)"
/note="Organ: blood; Vector: pTriplex2; Site_1: SfiI; Site_2: SfiI; A complementary DNA (cDNA) library from human erythroid precursor cells was constructed using SMART PCR (polymerase chain reaction) cDNA Library Construction Kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1 umol/L peptide nucleic acid (PNA) oligos (N-terminal)-biotin-GTC-CAC-CCG-AAG-CTT-G-(C-terminal) and (N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal). Synthesized cDNA was digested with SfiI and size-selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH Intramural Sequencing Center (NISC; Http://www.nisc.nih.gov/)."

```

ORIGIN

```

Query Match 32.3%; Score 405.6; DB 13; Length 587;
Best Local Similarity 89.8%; Pred. No. 1.1e-91;
Matches 465; Conservative 0; Mismatches 9; Indels 44; Gaps 1;

QY 1 ATGAGCTCTAAGTACCGCGGCTGTGCGCGCTGCTGCCCTCTGGGCCCTTAACACTG 60
Db |||
QY 70 ATGAGCTCTAAGTACCGCGGCTGTGCGCGCTGCTGCCCTCTGGGCCCTTAACACTG 129
Db |||
QY 61 GAAGCAGCTCTCATTTCTCTCTTCTATTTTACCACCTATGACGCTTCTTAGAGGAT 120
Db |||
QY 130 GAAGCAGCTCTCATTTCTCTCTTCTATTTTACCACCTATGACGCTTCTTAGAGGAT 189
QY 121 CAAAGGGGCTGTGGCATCTCTATC-----145
Db CAAAGGGGCTGTGGCATCTCTATCAAGGGGCTGGGAAGTCTGCATGCTCTATAAATC 249
QY 146 -----AGTTGCCAGATCTACCGTGTATGGGCCATCTGCTGGGCTTCTCA 196
Db 250 CAGAACGAGATCGGCCAGATCTGACGTTATGGCGGCCCTTGGGCTTCTCTCA 309
QY 197 CTTGAGTTTCCGAGACACAGCTGGAGCAGTGTGGCCCTTCAACCTCTTTCATGCTGGCG 256
Db 310 CTTCAATTTCCGAGACACAGCTGGAGCAGTGTGGCCCTTCAACCTCTTTCATGCTGGCG 369
QY 257 TTGGTGTGAGTGGGGAATCTGTCGACGCTTCTGACGCTTCTGACGCTTCTTGGGAGG 316
Db 370 TTGGTGTGAGTGGGGAATCTGTCGACGCTTCTGACGCTTCTTGGGAGG 429
QY 317 TGGTATACACTGTTTCAATTCGCTGTGCCACCATGATGCTTGTGCTGCTGCTATCT 376
Db 430 TGGTATACACTGTTTCAATTCGCTGTGCCACCATGATGCTTGTGCTGCTGCTATCT 489
QY 377 CAGTGGATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGTTGTTGATGTTGCTGGTG 436
Db 490 CAGCGGTGCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGTTGTTGATGTTGCTGGTG 549
QY 437 AGGTGACGCTTAGGCACCTGAGGATGTCATCAGT 474
Db 550 AGGTGACGCTTAGGCACCTGAGGATGTCATCAGT 587

```

RESULT 15

BUE59596

LOCUS

```

DEFINITION BUE59596 480 bp mRNA linear EST 30-SEP-2002
            cl49f09.z1 Hembase; Erythroid Precursor Cells (LCB:cl library) Homo
            sapiens cDNA clone cl49f09 5', mRNA sequence.

```

```

ACCESSION BUE59596
VERSION BUE59596.1 GI:23371778
EST.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

BUE59596
BUE59596.1 GI:23371778
EST.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Hembase; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 480)
Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
Gene Expression in Human Erythroid Precursor Cells
Unpublished (2002)
Contact: Jeffery L. Miller
Laboratory of Chemical Biology
National Institute of Diabetes and Digestive and Kidney Diseases
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
20892, USA
Tel: 301 402 2373
Fax: 301 435 5148
Email: jmf@nih.gov
The 'cl' library was constructed by Alexander Gubin, Ph.D. in the
Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
analyses by National Institutes of Health Intramural Sequencing
Center (NISC). More information available at:
http://hembase.nidk.nih.gov
Plate: 49 row: f column: 09
Seq primer: 5' lambda-triPLEX2 Sequencing Primer.
Location/Qualifiers
1..480
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="cl49f09"
/sex="unknown"
/tissue_type="blood"
/cell_type="Erythroid Precursor Cells"
/cell_line="Primary Culture of Peripheral Blood Mononuclear Cells"
/dev_stage="Precursor erythroblasts; GPA++"
/lab_host="DH5alpha"
/clone_lib="Hembase; Erythroid Precursor Cells (LCB:cl library)"
/note="Organ: blood; Vector: pTriplex2; Site_1: SfiI; Site_2: SfiI; A complementary DNA (cDNA) library from human erythroid precursor cells was constructed using SMART PCR (polymerase chain reaction) cDNA Library Construction Kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1 umol/L peptide nucleic acid (PNA) oligos (N-terminal)-biotin-GTC-CAC-CCG-AAG-CTT-G-(C-terminal) and (N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal). Synthesized cDNA was digested with SfiI and size-selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH Intramural Sequencing Center (NISC; Http://www.nisc.nih.gov/)."

```

```

FEATURES
source

```

ORIGIN

```

Query Match 31.8%; Score 399.2; DB 13; Length 480;
Best Local Similarity 98.1%; Pred. No. 4.1e-90;
Matches 404; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATGAGCTCTAAGTACCGCGGCTGTGCGCGGCTGCTGCCCTCTGGGCCCTTAACACTG 60
Db |||
QY 69 ATGAGCTCTAAGTACCGCGGCTGTGCGCGGCTGCTGCCCTCTGGGCCCTTAACACTG 128
Db |||
QY 61 GAAGCAGCTCTCATTTCTCTCTTCTATTTTACCACCTATGACGCTTCTTAGAGGAT 120
Db |||
QY 129 GAAGCAGCTCTCATTTCTCTCTTCTATTTTACCACCTATGACGCTTCTTAGAGGAT 188
QY 121 CAAAGGGGCTCTGTCGATCTCTATCAAGTTGGCCCAAGATCTGACGCTGATGGGCCATT 180

```

```

Db      189  CAAAAGGGCTCGTGGCATCCTATCAAGTCGGCCAGATCTGACCGTGATGGCGGCCCTT 248
Qy      181  GGCTTGGGCTTCCTCACCCTCGAGTTCCGAGAGACACAGCTGGAGCAGTGTGGCCTTCAAC 240
Db      249  GGCTTGGGCTTCCTCACCCTCAAAATTCGGAGAGACACAGCTGGAGCAGTGTGGCCTTCAAC 308
Qy      241  CTCTTCATGCTGGGCTTGGTGTGTCAGTGGGCAATCCTGCTGGACGGCTTCTCTGAGCCAG 300
Db      309  CTCTTCATGCTGGGCTTGGTGTGTCAGTGGGCAATCCTGCTGGACGGCTTCTCTGAGCCAG 368
Qy      301  TTCCTTCTGGGAAGTGTGTCATCACACTGTTCAAGTATTCGGCTGGGCCACCATGAGTGT 360
Db      369  TTCCTCCTGGGAAGTGTGTCATCACACTGTTCAAGTATTCGGCTGGGCCACCATGAGTGT 428
Qy      361  TTGTCGGTGTGATCTCAGTGGATGCTGCTTGGGGAGGTCAACTTGGGCGC 412
Db      429  ATGTGGTGTGATCTCAGCGGGTGTGTCTTGGGGAAGGTCAACTTGGGCGC 480

```

Search completed: August 25, 2004, 17:28:29
 Job time : 3123 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2004, 05:32:17 ; Search time 4907 Seconds
(without alignments)
11076.455 Million cell updates/sec

Title: US-09-600-714-41
Perfect score: 1254
Sequence: 1 atgaacttaagtaaccgcg.....atttgctgtggattttaa 1254

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl :

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg_hum.*

31: em.htg_inv.*

32: em.htg_other.*

33: em.htg_mus.*

34: em.htg_pln.*

35: em.htg_rnd.*

36: em.htg_mam.*

37: em.htg_vrt.*

38: em.sy.*

39: em.htgo_hum.*

40: em.htgo_mus.*

41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1254	100.0	1254	6	AX022514	AX022514 Sequence
2	1254	100.0	1254	6	BD124102	BD124102 Novel nuc
3	1252.4	99.9	1254	9	AB018969	AB018969 Homo sapi
4	1252.4	99.9	1354	9	HUMRHANT1	L08429 Human Rhd b
5	1252.4	99.9	1545	9	HSRHIII	X63094 H.sapiens m
6	1252.4	99.9	2790	9	HSRHXIII	X63097 H.sapiens m
7	1250.8	99.7	1254	9	AB018966	AB018966 Homo sapi
8	1250.8	99.7	1254	9	AB018967	AB018967 Homo sapi
9	1250.8	99.7	1480	9	AF510069	AF510069 Homo sapi
10	1250.8	99.7	1480	9	AF510070	AF510070 Homo sapi
11	1247.8	99.5	1251	9	AY449381	AY449381 Homo sapi
12	1247.8	99.5	1251	9	AY449382	AY449382 Homo sapi
13	1247.8	99.5	1251	9	AY449383	AY449383 Homo sapi
14	1247.8	99.5	1251	9	AY449384	AY449384 Homo sapi
15	1247.8	99.5	1251	9	AY449385	AY449385 Homo sapi
16	1247.6	99.5	1260	9	S78503	S78509 Rhd [human]
17	1246	99.4	1254	9	AB018968	AB018968 Homo sapi
18	1246	99.4	1254	9	S82449	S82449 Homo sapien
19	1242.8	99.1	1458	9	AF187846	AF187846 Homo sapi
20	1239.8	98.9	1251	9	S57971	S57971 Rh polypt
21	1239.6	98.9	1254	9	AB046420	AB046420 Homo sapi
22	1238	98.7	1328	9	AB049754	AB049754 Homo sapi
23	1231.4	98.2	1247	9	AF037626	AF037626 Homo sapi
24	1226.8	97.8	1254	9	AB037270	AB037270 Homo sapi
25	1225.2	97.7	1336	9	AF312679	AF312679 Homo sapi
26	1222	97.4	1330	9	HSZ97026	Z97026 Homo sapien
27	1220.4	97.3	1254	9	AF177938	AF177938 Homo sapi
28	1199.6	95.7	1328	9	AB030388	AB030388 Homo sapi
29	1199.6	95.7	1328	9	AB049753	AB049753 Homo sapi
30	1191.6	95.0	1350	9	GORRHB	L37053 Gorilla gor
31	1190	94.9	1430	9	AF510065	AF510065 Homo sapi
32	1188.4	94.8	1328	9	AB018644	AB018644 Homo sapi
33	1186.8	94.6	1350	9	CHPRHC	L37050 Pan troglod
34	1186.8	94.6	1430	9	AF510067	AF510067 Homo sapi
35	1186.8	94.6	1430	9	AF510068	AF510068 Homo sapi
36	1185.2	94.5	1384	6	I89819	I89819 Sequence 1
37	1185.2	94.5	1384	6	HUMRHPAA	M34015 Human Rh po
38	1185.2	94.5	1471	9	HSRH30A	X54534 Human mRNA
39	1183.6	94.4	1328	9	AB018645	AB018645 Homo sapi
40	1183.6	94.4	1350	9	CHPRHB	L37049 Pan troglod
41	1183.6	94.4	1430	9	AF510066	AF510066 Homo sapi
42	1182	94.3	1466	6	I89820	I89820 Sequence 2
43	1180.8	94.2	1427	9	AF056965	AF056965 Homo sapi
44	1180.6	94.1	1251	9	S57967	S57967 Rh polypt
45	1180.4	94.1	1350	9	CHPRHA	L37048 Pan troglod

ALIGNMENTS

RESULT 1
AX022514
LOCUS AX022514 1254 bp DNA linear PAT 24-NOV-2000
DEFINITION Sequence 41 from Patent WO9337763.
ACCESSION AX022514
VERSION AX022514.1 GI:10046112
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE
AUTHORS Flegel, W.A. and Wagner, F.F.
TITLE Novel nucleic acid molecules correlated with the rhesus weak d phenotype
JOURNAL Patent: WO 9337763-A 41 29-JUL-1999;

FLEGEL WILLY A (DE) ; WAGNER FRANZ F (DE) ; DRK BLUTSPENDEDIENST
BADEN WUE (DE)

FEATURES

source
1..1254
/location/Qualifiers
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
1..1254
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07879.1"
/db_xref="GI:10046113"
/db_xref="REMTREMBL:CAC07879"
/translation="MSKYPSPVRRCLPLWALTEALILLYFFHYDASLEDOXGL
VASYQVODITWAAAGLGLTSSPRHSWSVAFLNPLALGVQWAILLDGLSQQP
SGKVIITLIRIATLSALSVLSVDAVLKLVQVWLVVEVTAALGNLRVWISNI
FNTDYHNMHHIYFAAYFGLSVAMCLPKLPEDTQKQATIPSLSAMGALFLWM
FWPSPFALSRIEERKNVNTYAVAVSVTAISGSLAHQPKISKIYVHSVILA
GVAVGTSCHLIPSLAMVLGLVAGLISVGGAKLPCCNRVLGIPHSIMGNFSL
LGLLGEIITYVLVLVDITVAGNMGIMGFQVLLSIGELSLAIVIALTSGLLTGLLNLKI
WKAPHEAKYEDDDQVFWKFPFLAVGF"

ORIGIN

Query Match 100.0%; Score 1254; DB 6; Length 1254;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAGCTTAAGTACCGCGGTCTGTCGGCGGTGCTGCCCTCTCGGCGCTTAACTG 60
Db 1 ATGAGCTTAAGTACCGCGGTCTGTCGGCGGTGCTGCCCTCTCGGCGCTTAACTG 60
Qy 61 GAAGCAGCTCTCATCTCCCTCTCTATTTTACCCACATGATGAGCTTCTCTAGAGAT 120
Db 61 GAAGCAGCTCTCATCTCCCTCTCTATTTTACCCACATGATGAGCTTCTCTAGAGAT 120
Qy 121 CAAAGGGGCTCGTGGCATCTCAAGTTGGCCCAAGATCGACCGTGGCGGCATT 180
Db 121 CAAAGGGGCTCGTGGCATCTCAAGTTGGCCCAAGATCGACCGTGGCGGCATT 180
Qy 181 GGCTTGGGCTTCTCAGTCTGAGTTCCGAGACACAGCTGGAGCAGTGGCCCTCAAC 240
Db 181 GGCTTGGGCTTCTCAGTCTGAGTTCCGAGACACAGCTGGAGCAGTGGCCCTCAAC 240
Qy 241 CTCTTCATGTCGGCGTGTGTCAGTGGGCAATCTGCTGGAGCGCTTCTCGAGCGAG 300
Db 241 CTCTTCATGTCGGCGTGTGTCAGTGGGCAATCTGCTGGAGCGCTTCTCGAGCGAG 300
Qy 301 TTCCCTTCTGGGAGGTGGTCATCACACTGTTCAAGTATTCGGCTGGCCACCATGAGTGT 360
Db 301 TTCCCTTCTGGGAGGTGGTCATCACACTGTTCAAGTATTCGGCTGGCCACCATGAGTGT 360
Qy 361 TTGTCGCTGATCTCAGTGGATGCTCTTGGGAGGTCACTTGGCGAGTGGTGGT 420
Db 361 TTGTCGCTGATCTCAGTGGATGCTCTTGGGAGGTCACTTGGCGAGTGGTGGT 420
Qy 421 GTGATGTCGTGGTGGAGTGCACAGCTTTAGGCAACCTGAGATGTCATCAGTAATATC 480
Db 421 GTGATGTCGTGGTGGAGTGCACAGCTTTAGGCAACCTGAGATGTCATCAGTAATATC 480
Qy 481 TTCAACACAGATACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 481 TTCAACACAGATACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Qy 541 CTGTCGTGGCTGGTGGCTGGCAAGCTTCTACCGAGGAGGAGGAGGAGGAGGAGGAG 600
Db 541 CTGTCGTGGCTGGTGGCTGGCAAGCTTCTACCGAGGAGGAGGAGGAGGAGGAGGAG 600
Qy 601 ACAGCAACGATACCCAGTTTCTGTCGATGTCGTCGTCGTCGTCGTCGTCGTCGTCG 660
Db 601 ACAGCAACGATACCCAGTTTCTGTCGATGTCGTCGTCGTCGTCGTCGTCGTCGTCG 660
Qy 661 CCAAGTTTCACTCTGCTCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 661 CCAAGTTTCACTCTGCTCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 720

661 CCAAGTTTCACTCTGCTCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 720
721 ACCTACTATGCTCTAGCAGTACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 780
721 ACCTACTATGCTCTAGCAGTACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 780
781 CCCCAGGGAAGATCAGCAAGACTTATGTGCAAGTGGTGGTGGTGGTGGTGGTGGTGGT 840
781 CCCCAGGGAAGATCAGCAAGACTTATGTGCAAGTGGTGGTGGTGGTGGTGGTGGTGGT 840
841 GTGGGTACTCTGCTCACTGATCCCTTCTCCGTGGTGGTGGTGGTGGTGGTGGTGGT 900
841 GTGGGTACTCTGCTCACTGATCCCTTCTCCGTGGTGGTGGTGGTGGTGGTGGTGGT 900
901 GCTGGGTACTCTGCTCACTGATCCCTTCTCCGTGGTGGTGGTGGTGGTGGTGGTGGT 960
901 GCTGGGTACTCTGCTCACTGATCCCTTCTCCGTGGTGGTGGTGGTGGTGGTGGTGGT 960
961 GGGATTCCCCACAGTCCATCATGGGCTCACTGGGCTCACTGGGCTCACTGGGCTCACT 1020
961 GGGATTCCCCACAGTCCATCATGGGCTCACTGGGCTCACTGGGCTCACTGGGCTCACT 1020
1021 ATCATCTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
1021 ATCATCTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
1081 TTCAGTCTCTCTCAGCATTGGGAACTCAGCTTGGCACTGCTGCTGCTGCTGCTGCTGCT 1140
1081 TTCAGTCTCTCTCAGCATTGGGAACTCAGCTTGGCACTGCTGCTGCTGCTGCTGCTGCT 1140
1141 GGTCTCTCTCAGCATTGGGAACTCAGCTTGGCACTGCTGCTGCTGCTGCTGCTGCTGCT 1200
1141 GGTCTCTCTCAGCATTGGGAACTCAGCTTGGCACTGCTGCTGCTGCTGCTGCTGCTGCT 1200
1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTGGAAGTTTCTGGAAGTTTCTGGAAGTTT 1254
1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTGGAAGTTTCTGGAAGTTTCTGGAAGTTT 1254

RESULT 2

BD124102
LOCUS Novel nucleic acid molecule correlating to Rhesus weak D phenotype.
DEFINITION Novel nucleic acid molecule correlating to Rhesus weak D phenotype.
ACCESSION BD124102
VERSION BD124102.1 GI:23219047
KEYWORDS JP 2002500884-A/41.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1254)
AUTHORS Fregel, V.A. and Wagner, F.F.
TITLE Novel nucleic acid molecule correlating to Rhesus weak D phenotype
JOURNAL Patent: JP 2002500884-A 41 15-JAN-2002;
DRK BLUTSPENDEDIENST BADEN WUERTEMBERG GGBMB
COMMENT OS Unidentified
PN JP 2002500884-A/41
PD 15-JAN-2002
PF 18-DEC-1998 JP 2000528671
PR 23-JAN-1998 EP 98101203.2
PI VILLY A FREGEL, FRANZ F WAGNER
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12N5/00
PC C12P21/02, C12P21/08, C12Q1/02, C12Q1/68, G01N33/566, C12N15/00, PC
CC Strandedness: Single;
CC Topology: Linear;
CC Novel nucleic acid molecule correlating to Rhesus weak D CC
phenotype
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..1254
source


```
/organism="unidentified"  
/mol_type="genomic DNA"  
/db xref="taxon:32644"
```

ORIGIN

Query Match 100.0%; Score 1254; DB 6; Length 1254; Best Local Similarity 100.0%; Presd. No. 0; Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATGAGCTCTAAGTACCCGCGTCTCTCGGCGCTGCTGCCCTCTTGGGCCCTTAACACTG	60						
Db	1	ATGAGCTCTAAGTACCCGCGTCTCTCGGCGCTGCTGCCCTCTTGGGCCCTTAACACTG	60						
QY	61	GAAGCAGCTCTCAATCTCCTCTTCTAATTTTTTATCCCACTATGACGCTTCTCTTAGAGGAT	120						
Db	61	GAAGCAGCTCTCAATCTCCTCTTCTAATTTTTTATCCCACTATGACGCTTCTCTTAGAGGAT	120						
QY	121	CAAAAGGGGCTCGTGGCATCTTATCAAGTTGGCCAAAGATCTGACCGTATGCGCGGCATT	180						
Db	121	CAAAAGGGGCTCGTGGCATCTTATCAAGTTGGCCAAAGATCTGACCGTATGCGCGGCATT	180						
QY	181	GGCTTGGGCTTCCCTCACCTCGAGTTTCGGAGACACAGCTGGAGCAGTGGCCCTTCAAC	240						
Db	181	GGCTTGGGCTTCCCTCACCTCGAGTTTCGGAGACACAGCTGGAGCAGTGGCCCTTCAAC	240						
QY	241	CTCTTTCATGCTGGCGCTTGCTGTCAGTGGGCAATCCTGCTGGACGGCTCTCTGAGCCAG	300						
Db	241	CTCTTTCATGCTGGCGCTTGCTGTCAGTGGGCAATCCTGCTGGACGGCTCTCTGAGCCAG	300						
QY	301	TTCCCTTCTGGGAGGTGCTCATCACACTGTTTCAGTATTCGGCTGGGCCACCATGAGTGC	360						
Db	301	TTCCCTTCTGGGAGGTGCTCATCACACTGTTTCAGTATTCGGCTGGGCCACCATGAGTGC	360						
QY	361	TTGTTCGCTGCTGATCTCAGTGGATGCTGTTTGGGGAAGGTCAACTTGGCGCAGTTCGGTG	420						
Db	361	TTGTTCGCTGCTGATCTCAGTGGATGCTGTTTGGGGAAGGTCAACTTGGCGCAGTTCGGTG	420						
QY	421	GTGATGCTGCTGGTGGAGGTGACAGCTTTTAGGCAACCTGAGGATGGTCAATCAATATC	480						
Db	421	GTGATGCTGCTGGTGGAGGTGACAGCTTTTAGGCAACCTGAGGATGGTCAATCAATATC	480						
QY	481	TTCAACACAGACTACCAATGAACTATCATGATGACATCTACCTGTTCCGAGCCTATTTGGG	540						
Db	481	TTCAACACAGACTACCAATGAACTATGATGACATCTACCTGTTCCGAGCCTATTTGGG	540						
QY	541	CTGTCTGTGGCTGGTGCCTGCCAAAGCCCTTACCCGAGGGAACGGAGGATAAAGATCAG	600						
Db	541	CTGTCTGTGGCTGGTGCCTGCCAAAGCCCTTACCCGAGGGAACGGAGGATAAAGATCAG	600						
QY	601	ACAGCAACGATACCCAGTTTGTCTGCCATGCTGGGGCCCTCTTCTTGTGGAATGTTCTGG	660						
Db	601	ACAGCAACGATACCCAGTTTGTCTGCCATGCTGGGGCCCTCTTCTTGTGGAATGTTCTGG	660						
QY	661	CCAAGTTTCAACTCTGCTCTGCTGAGAAGTCCAAATCGAAAGGAAGATGCGGTGTTCAAC	720						
Db	661	CCAAGTTTCAACTCTGCTCTGCTGAGAAGTCCAAATCGAAAGGAAGATGCGGTGTTCAAC	720						
QY	721	ACCTACTATGCTGTAGCAGTACAGGTGGTGACAGCCATCTCAGGGTCAATCCTTTGGCTCAC	780						
Db	721	ACCTACTATGCTGTAGCAGTACAGGTGGTGACAGCCATCTCAGGGTCAATCCTTTGGCTCAC	780						
QY	781	CCCCAAGGGAAGATCAGCAAGCTTATGTGCAAGTGCAGTGCCTGTTGGCAGAGCGGTGGCT	840						
Db	781	CCCCAAGGGAAGATCAGCAAGCTTATGTGCAAGTGCAGTGCCTGTTGGCAGAGCGGTGGCT	840						
QY	841	GTGGGTACCTCGTGTCACTTGATCCCTTCTCCGTGGCTTGGCATGCTGCTGGGTCTTGTG	900						
Db	841	GTGGGTACCTCGTGTCACTTGATCCCTTCTCCGTGGCTTGGCATGCTGCTGGGTCTTGTG	900						
QY	901	GCTTGGGCTGATCTCCGTCGGGGAGCCAAAGTACTCTGCGGGGTGTTGTAAACCGAGTGC	960						
Db	901	GCTTGGGCTGATCTCCGTCGGGGAGCCAAAGTACTCTGCGGGGTGTTGTAAACCGAGTGC	960						

Qy	961	GGGATTC	CCCA	CAGCTCC	ATCATGGG	CTCA	CAACTT	CAGCTT	GTGTGG	GTCTG	CTTGG	AGAG	1020
Db	961	GGGATTC	CCCA	CAGCTCC	ATCATGGG	CTCA	CAACTT	CAGCTT	GTGTGG	GTCTG	CTTGG	AGAG	1020
Qy	1021	ATCATCT	ACAT	TGCTG	TGGTGC	TTGAT	TACCGT	CGGAGC	CGGC	CAATGC	ATCAT	TGGC	1080
Db	1021	ATCATCT	ACAT	TGCTG	TGGTGC	TTGAT	TACCGT	CGGAGC	CGGC	CAATGC	ATCAT	TGGC	1080
Qy	1081	TTCAGGT	CTCT	CTCAG	CATTGG	GGGA	ACTCAG	CTTGG	CCCAT	CTGTAG	CTCTC	CACGTCT	1140
Db	1081	TTCAGGT	CTCT	CTCAG	CATTGG	GGGA	ACTCAG	CTTGG	CCCAT	CTGTAG	CTCTC	CACGTCT	1140
Qy	1141	GGTCTC	CTGAC	AGTTG	CTCTTA	AACTT	TAAAT	TGGA	AAC	CACTCAT	GAGG	CTAAA	1200
Db	1141	GGTCTC	CTGAC	AGTTG	CTCTTA	AACTT	TAAAT	TGGA	AAC	CACTCAT	GAGG	CTAAA	1200
Qy	1201	TATTTG	ATGAC	CAAG	TTTTCT	GAA	AGTTT	CTC	CA	TGGCT	TTG	ATTTTAA	1254
Db	1201	TATTTG	ATGAC	CAAG	TTTTCT	GAA	AGTTT	CTC	CA	TGGCT	TTG	ATTTTAA	1254

RESULT 3

AB018969

LOCUS

DEFINITION

AB018969 Homo sapiens RhdY0 mRNA for Rh blood group D antigen (RHD), complete cds.

1254 bp

linear

PRI 03-JUL-1999

RESULT 3

RESULTS 3	AB018969	1254 bp	linear	PRI 03-JUL-1999
LOCUS	AB018969	1254 bp	linear	PRI 03-JUL-1999
DEFINITION	AB018969	1254 bp	linear	PRI 03-JUL-1999
DESCRIPTION	AB018969	1254 bp	linear	PRI 03-JUL-1999
ACCESSION	AB018969	1254 bp	linear	PRI 03-JUL-1999
VERSION	AB018969	1254 bp	linear	PRI 03-JUL-1999
KEYWORDS	AB018969	1254 bp	linear	PRI 03-JUL-1999
SOURCE	AB018969	1254 bp	linear	PRI 03-JUL-1999
ORGANISM	AB018969	1254 bp	linear	PRI 03-JUL-1999
REFERENCE	AB018969	1254 bp	linear	PRI 03-JUL-1999
AUTHORS	AB018969	1254 bp	linear	PRI 03-JUL-1999
TITLE	AB018969	1254 bp	linear	PRI 03-JUL-1999
JOURNAL	AB018969	1254 bp	linear	PRI 03-JUL-1999
REFERENCE	AB018969	1254 bp	linear	PRI 03-JUL-1999
AUTHORS	AB018969	1254 bp	linear	PRI 03-JUL-1999
TITLE	AB018969	1254 bp	linear	PRI 03-JUL-1999
JOURNAL	AB018969	1254 bp	linear	PRI 03-JUL-1999

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>...</i>	
5. <i>...</i>	
6. <i>...</i>	
7. <i>...</i>	
8. <i>...</i>	
9. <i>...</i>	
10. <i>...</i>	
11. <i>...</i>	
12. <i>...</i>	
13. <i>...</i>	
14. <i>...</i>	
15. <i>...</i>	
16. <i>...</i>	
17. <i>...</i>	
18. <i>...</i>	
19. <i>...</i>	
20. <i>...</i>	
21. <i>...</i>	
22. <i>...</i>	
23. <i>...</i>	
24. <i>...</i>	
25. <i>...</i>	
26. <i>...</i>	
27. <i>...</i>	
28. <i>...</i>	
29. <i>...</i>	
30. <i>...</i>	
31. <i>...</i>	
32. <i>...</i>	
33. <i>...</i>	
34. <i>...</i>	
35. <i>...</i>	
36. <i>...</i>	
37. <i>...</i>	
38. <i>...</i>	
39. <i>...</i>	
40. <i>...</i>	
41. <i>...</i>	
42. <i>...</i>	
43. <i>...</i>	
44. <i>...</i>	
45. <i>...</i>	
46. <i>...</i>	
47. <i>...</i>	
48. <i>...</i>	
49. <i>...</i>	
50. <i>...</i>	
51. <i>...</i>	
52. <i>...</i>	
53. <i>...</i>	
54. <i>...</i>	
55. <i>...</i>	
56. <i>...</i>	
57. <i>...</i>	
58. <i>...</i>	
59. <i>...</i>	
60. <i>...</i>	
61. <i>...</i>	
62. <i>...</i>	
63. <i>...</i>	
64. <i>...</i>	
65. <i>...</i>	
66. <i>...</i>	
67. <i>...</i>	
68. <i>...</i>	
69. <i>...</i>	
70. <i>...</i>	
71. <i>...</i>	
72. <i>...</i>	
73. <i>...</i>	
74. <i>...</i>	
75. <i>...</i>	
76. <i>...</i>	
77. <i>...</i>	
78. <i>...</i>	
79. <i>...</i>	
80. <i>...</i>	
81. <i>...</i>	
82. <i>...</i>	
83. <i>...</i>	
84. <i>...</i>	
85. <i>...</i>	
86. <i>...</i>	
87. <i>...</i>	
88. <i>...</i>	
89. <i>...</i>	
90. <i>...</i>	
91. <i>...</i>	
92. <i>...</i>	
93. <i>...</i>	
94. <i>...</i>	
95. <i>...</i>	
96. <i>...</i>	
97. <i>...</i>	
98. <i>...</i>	
99. <i>...</i>	
100. <i>...</i>	

```

1. 1254
/organism="Homo sapiens"
/mol_type="mRNA"
/isolate="YO"
/db_xref="taxon:9606"
1. 1254
/gene="RhDYO"
1. 1254
/gene="RhDYO"
/note="RhD variant RhDYO"
/codon_start=1
/product="Rh blood group D antigen (RHD) "
/protein_id="BAA82159.1"
/db_xref="GI:5360899"
/translation="MSSKYPKSRVRCFLPMALTLEAAILLLFYPTHTYDASIEDOKGL
VASVYQGDGLTWMAIGFELTSSPRFSSSVAFNPLMALGQVMAILLDGFYSQFP
SGKVYITLFTSLATMSALSVLISVDVAIGKVNLAQLVNVLVEVTAIGNLRWVLSNI
FNIDTHNNMHIIYFAAIFGLSVACWLPEKPIPEGTDKDTATPSLSAMIGALFLIM
FWPFSNATLSRSPKIRKNAVENTYVAVSVVTAISGSLAHPOQKI SKTYVHGAVLA
GGVATGTCILIPSPWLAMVLIGSLVSGVAKVILPGCCNRVLGIPHSSITMGYNFSL
LLGLGEIIYVILLVLTGAGNGMIGFQWLLSIGELSLAIVALTSGLLTGLLLNLKI
WKAPAEAKYFDDDFDPEWPKTFLHVAVG"

```

ORIGIN

Query Match 99.9%; Score 1252.4; DB 9; Length 1254;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ATGAGCTCTAAGTACCGCGGTCGTCCCGCGTGTGCTGCCCTCTGTGGGCCCTTAAACATG	60
Db	1	ATGAGCTCTAAGTACCGCGGTCGTCCCGCGTGTGCTGCCCTCTGTGGGCCCTTAAACATG	60
QY	61	GAAGCAGCTCTCATCTCCCTCTCTATTTTACCCACTATGACGCTTCCCTTAGAGGAT	120
Db	61	GAAGCAGCTCTCATCTCCCTCTCTATTTTACCCACTATGACGCTTCCCTTAGAGGAT	120
QY	121	CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAAAGATCTGACCGGTGATGGCGGCATT	180
Db	121	CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAAAGATCTGACCGGTGATGGCGGCATT	180
QY	181	GGCTTGGCTTCTCTACCTCGAGTTTCCGGAGACACAGCTGGAGCATGTGGCTTCAAC	240
Db	181	GGCTTGGCTTCTCTACCTCGAGTTTCCGGAGACACAGCTGGAGCATGTGGCTTCAAC	240
QY	241	CTCTTCATGCTGGCGTGTGGTGGCAATCTCTGCGGCAATCTCTGAGCGCTTCTTGAGCCAG	300
Db	241	CTCTTCATGCTGGCGTGTGGTGGCAATCTCTGCGGCAATCTCTGAGCGCTTCTTGAGCCAG	300
QY	301	TTCCCTTCTGGAGAGTGGTCATCACACTGTTTCAGTATTCGGCTGGCCACCATGAGTCT	360
Db	301	TTCCCTTCTGGAGAGTGGTCATCACACTGTTTCAGTATTCGGCTGGCCACCATGAGTCT	360
QY	361	TTGTCGGTGTGATCTCAGTGGATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGGTG	420
Db	361	TTGTCGGTGTGATCTCAGTGGATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGGTG	420
QY	421	GTGATGTGTGTGGAGTGCACAGCTTTAGGCAACCTGAGATGTGTCACTAGTAATATC	480
Db	421	GTGATGTGTGTGGAGTGCACAGCTTTAGGCAACCTGAGATGTGTCACTAGTAATATC	480
QY	481	TTCAACACAGACTACCATGAACATGATGACATCTACGTGTTCGAGGCTATTTGGG	540
Db	481	TTCAACACAGACTACCATGAACATGATGACATCTACGTGTTCGAGGCTATTTGGG	540
QY	541	CTGTCTGTGGCTGGTGCCTCCAAAGCTTACCGAGGAAACGGAGGATAAGATCAG	600
Db	541	CTGTCTGTGGCTGGTGCCTCCAAAGCTTACCGAGGAAACGGAGGATAAGATCAG	600
QY	601	ACAGCAAGTACCCAGTTCCTCTGCCATGTGGCGCCCTCTTCTTGTGGATGTCTTG	660
Db	601	ACAGCAAGTACCCAGTTCCTCTGCCATGTGGCGCCCTCTTCTTGTGGATGTCTTG	660
QY	661	CCAAAGTTTCAACTCTGTCTCTGAGAGTCCAATCGAAAGAAATGCCGTGTTCAC	720
Db	661	CCAAAGTTTCAACTCTGTCTCTGAGAGTCCAATCGAAAGAAATGCCGTGTTCAC	720
QY	721	ACCTACTATGCTGTAGCAGTCCAGTGTGACAGCCATCTCAGGCTCATCTTGGCTCAC	780
Db	721	ACCTACTATGCTGTAGCAGTCCAGTGTGACAGCCATCTCAGGCTCATCTTGGCTCAC	780
QY	781	CCCCAAGGGAAGATCAGCAAGATTAATGTGCACAGTGGCGTGTGGCAGGCGGTGCT	840
Db	781	CCCCAAGGGAAGATCAGCAAGATTAATGTGCACAGTGGCGTGTGGCAGGCGGTGCT	840
QY	841	GTGGGTACCTCGTGTCACTGATCCCTTCTCCGTGGCTTGGCATGGTCTGGCTTGTG	900
Db	841	GTGGGTACCTCGTGTCACTGATCCCTTCTCCGTGGCTTGGCATGGTCTGGCTTGTG	900
QY	901	GCTGGCTGATCTCGTGGGGAGCCAAAGTACCTTCCGGGGTGTGTAAACCGAGTGTG	960
Db	901	GCTGGCTGATCTCGTGGGGAGCCAAAGTACCTTCCGGGGTGTGTAAACCGAGTGTG	960
QY	961	GGGATTTCCCAAGCTCCATCATGGGCTTCAACTTTTACCTTCTGGGTCTGCTGGAG	1020
Db	961	GGGATTTCCCAAGCTCCATCATGGGCTTCAACTTTTACCTTCTGGGTCTGCTGGAG	1020
QY	1021	ATCATCTACATGTGCTGTGTGATACCGTTCGGAGCGGCAATGGCATGATTCG	1080
Db	1021	ATCATCTACATGTGCTGTGTGATACCGTTCGGAGCGGCAATGGCATGATTCG	1080
QY	1081	TTCCAGGCTCTCTCAGCATTTGGGAACCTGAGCTTGGCCATCTGATAGCTCTCACGCT	1140
Db	1081	TTCCAGGCTCTCTCAGCATTTGGGAACCTGAGCTTGGCCATCTGATAGCTCTCACGCT	1140
QY	1141	GGTCTCTCAGCAGTTTGTCTCTAAATCTTAAATATGAAACACCTCATGAGGCTAAA	1200
Db	1141	GGTCTCTCAGCAGTTTGTCTCTAAATCTTAAATATGAAACACCTCATGAGGCTAAA	1200
QY	1201	TATTTTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTTTGATTTTAA	1254
Db	1201	TATTTTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTTTGATTTTAA	1254
RESULT 4			
HUMRHDAINTI			
LOCUS			
DEFINITION			
HUMAN RHD blood group antigen mRNA, complete cds.			
ACCESSION			
L08429			
VERSION			
L08429.1 GI:37390			
KEYWORDS			
Rh blood group; Rhd blood group; antigen; blood group antigen.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE			
1 (Bases 1 to 1354)			
AUTHORS			
Arce, M.A., Thompson, E.S., Wagner, S., Coyne, K.E., Ferdman, B.A. and			
Lublin, D.M.			
TITLE			
Molecular cloning of Rhd cDNA derived from a gene present in			
RHD-positive, but not RHD-negative individuals			
JOURNAL			
Blood 82 (2), 651-655 (1993)			
MEDLINE			
93320449			
PUBMED			
8329718			
COMMENT			
Original source text: Homo sapiens (library: HL1058b (from			
Clontech)) Bone marrow cDNA to mRNA.			
FEATURES			
source			
1..1354			
Location/Qualifiers			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
1..1254			
/standard_name="Rhd"			
/citations=[1]			
/codon_start=1			
/evidence=experimental			
/product="Rhd blood group antigen"			
/protein_id="AA02679.1"			
/db_xref="GI:37391"			
/translation="MSSKPRVRRCLPLWALTLEALILLFFTHYDASLEQKGL			
VASYQVQDLTVMAAIGLFTSPRRHSWVAFNLFMLALGVQWAILLDGFIQSP			
SGKVIITLFTIRLATLSVDAVLGKVNLAQLVVMVLVEVTLGNLRLMVISNI			
FNTDYHMNMWHIVFAAYFGLSVAMCLPKPEGTEDKDTATIPSLAMLGALFLWM			
FWPFSNALLRSPIERKNVNTYVAVSVVTAISGSLAHPOGKISKTVHSVAVLA			
GGVAVTSCHLIPSPMLAVGLVAGLISVGGAKYLPCCNRVLTGIPHSIMGVNFSL			
LGLLGEIYIIVLLDVTGAGNGMIGFQVLLSISIGLSLAIVALTSGLLTGLLLNLKI			
WKAPHEAKYFDDQVFWKPHLVAGF"			
ORIGIN			
Query Match			
99.9%; Score 1252.4; DB 9; Length 1354;			
Best Local Similarity			
99.9%; Pred. No. 0;			
Matches 1253; Conservative			
0; Mismatches			
1; Indels			
0; Gaps			
QY	1	ATGAGCTCTAAGTACCGCGGTCGTCCCGCGTGTGCTGCCCTCTGTGGGCCCTTAAACATG	60
Db	1	ATGAGCTCTAAGTACCGCGGTCGTCCCGCGTGTGCTGCCCTCTGTGGGCCCTTAAACATG	60
QY	61	GAAGCAGCTCTCATCTCCCTCTCTATTTTACCCACTATGACGCTTCCCTTAGAGGAT	120
Db	61	GAAGCAGCTCTCATCTCCCTCTCTATTTTACCCACTATGACGCTTCCCTTAGAGGAT	120
QY	121	CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAAAGATCTGACCGGTGATGGCGGCATT	180
Db	121	CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAAAGATCTGACCGGTGATGGCGGCATT	180

Qy	121	CAAAAGGGGCTCGTGGATCTCTATCAAGTTTGGCCAAGATCTGACCGTGAATGGCGGCATT	180
Db	179	CAAAAGGGGCTCGTGGATCTCTATCAAGTTTGGCCAAGATCTGACCGTGAATGGCGGCATT	238
Qy	181	GGCTTCGGCTTCCTCACCTCGAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCTCAAC	240
Db	239	GGCTTCGGCTTCCTCACCTCGAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCTCAAC	298
Qy	241	CTCTTCATGCTGCGCTTGGTGTGCAGTGGGCAATCCTGCTGGAACGGCTTCCTGAGCCAG	300
Db	299	CTCTTCATGCTGCGCTTGGTGTGCAGTGGGCAATCCTGCTGGAACGGCTTCCTGAGCCAG	358
Qy	301	TTCCCTCTCTGGGAAGTGGTCAACACACTGTTTCAGTATTTCGGCTGCGCCACCATGAGTGTCT	360
Db	359	TTCCCTCTCTGGGAAGTGGTCAACACACTGTTTCAGTATTTCGGCTGCGCCACCATGAGTGTCT	418
Qy	361	TTGTGCGGTGCTGATCTCAGTGGATGCTGTTCTTGGGAAGTCAACTTGGCGCAGTTGGTG	420
Db	419	TTGTGCGGTGCTGATCTCAGTGGATGCTGTTCTTGGGAAGTCAACTTGGCGCAGTTGGTG	478
Qy	421	GTGATGGTGTGGTGGAGGTGAAGCTTTTAGGCAACCTGAGGATGCTCATCAGTAAATATC	480
Db	479	GTGATGGTGTGGTGGAGGTGAAGCTTTTAGGCAACCTGAGGATGCTCATCAGTAAATATC	538
Qy	481	TTCAACACACACTACCACATGAACATGATGCACATCTACGTGTTTCGACGCTATTTTGGG	540
Db	539	TTCAACACACACTACCACATGAACATGATGCACATCTACGTGTTTCGACGCTATTTTGGG	598
Qy	541	CTGTCTGTGCGTGGTGCCTCCGCAAGCCCTCTACCGAGGGAACGGAGGATAAAGATCAG	600
Db	599	CTGTCTGTGCGTGGTGCCTCCGCAAGCCCTCTACCGAGGGAACGGAGGATAAAGATCAG	658
Qy	601	ACAGCACAGTACCCAGTTTGTCTGCCATGCTGCGGCGCCTCTTCTTGTGGAATGTTCTGG	660
Db	659	ACAGCACAGTACCCAGTTTGTCTGCCATGCTGCGGCGCCTCTTCTTGTGGAATGTTCTGG	718
Qy	661	CCAAGTTTCAACTCTGCTCTGCTGAGAGTCCAAATCGAAAGGAAGTCCGCTGTTTCAAC	720
Db	719	CCAAGTTTCAACTCTGCTCTGCTGAGAGTCCAAATCGAAAGGAAGTCCGCTGTTTCAAC	778
Qy	721	ACTTACTATGCTGTAGCAGTCAAGGTGGTGCACAGCCATCTCAGGTCATCCTTGGCTCAC	780
Db	779	ACTTACTATGCTGTAGCAGTCAAGGTGGTGCACAGCCATCTCAGGTCATCCTTGGCTCAC	838
Qy	781	CCCCAAGGGAAGATCAGCAAGAATTAATGTGCACAGTCCGCTGTTGGCAGAGGCGTGGCT	840
Db	839	CCCCAAGGGAAGATCAGCAAGAATTAATGTGCACAGTCCGCTGTTGGCAGAGGCGTGGCT	898
Qy	841	GTGGGTACTGCTGTACCTGATCCCTTCTCCGTGGCTTGCCATGCTGGTCTGTGCTG	900
Db	899	GTGGGTACTGCTGTACCTGATCCCTTCTCCGTGGCTTGCCATGCTGGTCTGTGCTG	958
Qy	901	GCTGGGCTGATCTCCGTCCGGGGAGCCAAGTACTCGCCGGGTGTTCTTAACCCAGTGCCTG	960
Db	959	GCTGGGCTGATCTCCGTCCGGGGAGCCAAGTACTCGCCGGGTGTTCTTAACCCAGTGCCTG	1018
Qy	961	GGGATTCGCCACAGCTCATTATGGGCTAACATTCACGTTGCTGGGTCTGCTTGGAGAG	1020
Db	1019	GGGATTCGCCACAGCTCATTATGGGCTAACATTCACGTTGCTGGGTCTGCTTGGAGAG	1078
Qy	1021	ATCATCTACATTTGCTGCTGGTCTTGTATACCGTCCGAGCCGCAATGGCATGATTGGC	1080
Db	1079	ATCATCTACATTTGCTGCTGGTCTTGTATACCGTCCGAGCCGCAATGGCATGATTGGC	1138
Qy	1081	TTCCAGGTCTCTCCTCAGCATTTGGGAACTCAGCTTGGCCATCGTGATAGCTCTCAAGTCT	1140
Db	1139	TTCCAGGTCTCTCCTCAGCATTTGGGAACTCAGCTTGGCCATCGTGATAGCTCTCAAGTCT	1198
Qy	1141	GGTCTCTGACAGGTTTGTCTCTTAATCTTAAATATGGAACACACCTCATGAGCTAAA	1200
Db	1199	GGTCTCTGACAGGTTTGTCTCTTAATCTTAAATATGGAACACACCTCATGAGCTAAA	1258
Qy	1201	TATTTTGTATGATCAACAGTTTCTCTGGAAGTTTCCCTCATTTGGCTGTGTGATTTTAA	1254

Db 1259 TATTTTGTACCAAGTTTCTGGAAGTTCCTCATTTGGCTGTGGAATTTAA 1312

RESULT 6
 HSRHXIII
 LOCUS 2790 bp mRNA linear PRI 07-JUL-1997
 DEFINITION H.sapiens mRNA for rhesus polypeptide (RH XIII).
 ACCESSION X63097.1 S48953
 VERSION X63097.1 GI:36046
 KEYWORDS isoform; red cell membrane protein; Rh blood group; Rh polypeptide;
 rhesus polypeptide.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2790)
 AUTHORS Le Van Kim,C., Cherif-Zahar,B., Raynal,V., Mouro,I., Lopez,M.,
 Carton,J.P. and Colin,Y.
 TITLE Multiple Rh messenger RNA isoforms are produced by alternative
 splicing
 JOURNAL Blood 80 (4), 1074-1078 (1992)
 MEDLINE 92360855
 PUBMED 1379850
 REFERENCE 2
 AUTHORS Le van Kim,C., Mouro,I., Cherif-Zahar,B., Raynal,V., Cherrier,C.,
 Carton,J.P. and Colin,Y.
 TITLE Molecular cloning and primary structure of the human blood group
 RHD polypeptide
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89 (22), 10925-10929 (1992)
 MEDLINE 92066356
 PUBMED 1438298
 REFERENCE 3 (bases 1 to 2790)
 AUTHORS Colin,Y.
 TITLE Direct Submission
 JOURNAL Submitted (04-NOV-1991) Y. Colin, INSERM, U76, INTS, 6, Rue
 A.Cabanef, 75015 Paris, FRANCE
 COMMENT See also X63094-98 & M34015.
 FEATURES Location/Qualifiers
 1..2790
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1p35"
 /clone="RH XIII"
 /tissue_type="bone marrow"
 /tissue_lib="human bone marrow cDNA"
 17..1270
 /codon_start=1
 /product="Rhesus polypeptide XIII"
 /protein_id="CAA44811.1"
 /db_xref="GI:36047"
 /db_xref="GOA:Q02161"
 /db_xref="SWISS-PROT:Q02161"
 /translation="MSKYPFRVRCLPLRWALTLEALILLFFYFTHVDASLEDKGL
 VASQVQDUTVAAIGSLGSLTSPFRHWSVAFNLFMLALGWQALILDGFLSQFP
 SKGVITLFSIRLATAGSLVSLISVDALVKVNAQLVNVVEVTALGNLRWVLSIN
 FNTDYNNMMHIIYFAAYFGLSVAMCLPELPEGTEDKQDTATIPSLASMLGALFLMI
 FWPSPNSALRSPIERKNVFNFTYYAVASVTAISGSSLAHPQGKISKTYVHSAFLA
 GGVAVTGSCHLIPSPMLAMVLGI;VAGLIISVGAKYILPGCCNRVLGIPHSSIMGVNFSL
 LGLLGLIYIVLVLDTVGAGNGMIGFQVLLSIGLSLAIVALTSGLLTGILLNLNLKI
 WKPAHEAKYFDQVFKFPHLVGF"

CDS
 99.9%; Score 1252.4; DB 9; Length 2790;
 Query Match 99.9%;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAGCTCTAAGTACC CGCGTCTGTGTCGGCGTCTGCTGCCCTCTTGGGCCCTAACACTG 60
 Db 17 ATGAGCTCTAAGTACC CGCGTCTGTGTCGGCGTCTGCTGCCCTCTTGGGCCCTAACACTG 76

RESULT 8	AB018967	1254 bp	mRNA	linear	PRI 03-JUL-1999
LOCUS	Homo sapiens RhdVa (TO)	mRNA for Rh blood group D antigen (RHD)			
DEFINITION	complete cds.				
ACCESSION	AB018967	GI:5360244			
VERSION	RHDVa (TO); Rh blood group D antigen (RHD)				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (sites)				
AUTHORS	Hyodo,H., Ishikawa,Y., Kashiwase,K., Ogawa,A., Watanabe,Y., Tsunoyama,H., Toyoda,C., Uchikawa,M., Akaza,T. and Fujii,T.				
TITLE	Polymorphisms of RhdVa in Japanese				
REFERENCE	2 (bases 1 to 1254)				
AUTHORS	Uchikawa,M., Hyodo,H. and Ishikawa,Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-OCT-1998) Hironobu Hyodo, Japanese Red Cross Central Blood Center, Research; 4-1-31, Hiroo, Shibuya, Tokyo 150-0012, Japan (E-mail:hyodo@hla.cbc.jrc.or.jp, Tel:+81-3-5485-6009, Fax:+81-3-3406-7892)				
FEATURES	Location/Qualifiers				
source	1..1254	/organism="Homo sapiens"			
		/mol_type="mRNA"			
		/isolate="TO"			
gene	1..1254	/db_xref="taxon:9606"			
		/genes="RhdVa (TO)"			
CDS	1..1254	/gene="RhdVa (TO)"			
		/note="RhD variant RhdVa (TO)"			
		/codon_start=1			
		/product="Rh blood group D antigen (RHD)"			
		/protein_id="BA081900.1"			
		/db_xref="GI:5360245"			
		/translation="MSSKPSRVRRCLPLWALTLEAAILLFFVFFTHYDASLEBQKGLVASYQVQDLTVMAAIGLFTSTFRHSWSVAFLNFMALGQVMAILLDFGQSOPFSGKVVILFSLRLATMSALSVLISVDALGKVNLAQLVVMVLVEVTLGNLRMVISINPNTDYHNMWHIYVFAAYFGLSVAMCLPKLPETEDKDTATIPSLAMGLALFLMWWFWSFNALLRSPIQKXNAMPNTYYAVASVVTALSGSLAHQPKISKTYVHSAVLAGGVAVGSCHLIPSLMIAVLGLVAGLISVGAKYLPCCNRVLGIPIHSSIMGNFSLVIGLGEIYIYLVLLDTVGAGNGMIGFQVLLISIGELSLAIVIALTSGLLTGLLLNLKIKWAPHEAKYFDDQVFWKFPFLAVGF"			
ORIGIN					
Query Match	99.7%;	Score 1250.8;	DB 9;	Length 1254;	
Best Local Similarity	99.8%;	Pred. No. 0;			
Matches 1252;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;	
QY	1	ATGAGCTCTAAGTACCGCGGTCTGTCCGGCGCTGCCTGCCCTCTGGGCCCTTAACACTG	60		
Db	1	ATGAGCTCTAAGTACCGCGGTCTGTCCGGCGCTGCCTGCCCTCTGGGCCCTTAACACTG	60		
QY	61	GAAGCAGCTCTCAATCTCTCTTCTATTTTATTTTACCACTATGACGCTTCCTTTAGGAT	120		
Db	61	GAAGCAGCTCTCAATCTCTCTCTTCTATTTTATTTTACCACTATGACGCTTCCTTTAGGAT	120		
QY	121	CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCCAAGATCTGACCGTGATGGCGGCCATT	180		
Db	121	CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCCAAGATCTGACCGTGATGGCGGCCATT	180		
QY	181	GGCTTGGGCTTCCCTCACCTCGAGTTTTCGGAGACACAGCTGGAGCAGTGTGGCCTTCAAC	240		
Db	181	GGCTTGGGCTTCCCTCACCTCGAGTTTTCGGAGACACAGCTGGAGCAGTGTGGCCTTCAAC	240		
QY	241	CTCTTCATGCTGGCGCTTGTGTGCAAGTGGGCAATCTCTGCTGGAGCGGCTTCTTGAGCCAG	300		
Db	241	CTCTTCATGCTGGCGCTTGTGTGCAAGTGGGCAATCTCTGCTGGAGCGGCTTCTTGAGCCAG	300		


```
QY 301 TTCCCTTCTGGAGGTGGTCAATCACACTGTTTCACTATTCGGCTGGCCACCATGAGTCT 360
Db 301 TTCCCTTCTGGAGGTGGTCAATCACACTGTTTCACTATTCGGCTGGCCACCATGAGTCT 360
QY 361 TTGTGGTGTGATCTCAGTGGATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTGGTG 420
Db 361 TTGTGGTGTGATCTCAGTGGATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTGGTG 420
QY 421 GTGATGGTGTGGAGGTGACAGCTTTAGGCAACTGAGGATGTCATCAGTAATATC 480
Db 421 GTGATGGTGTGGAGGTGACAGCTTTAGGCAACTGAGGATGTCATCAGTAATATC 480
QY 481 TTCAACACAGACTACCAATGAATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 481 TTCAACACAGACTACCAATGAATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 541 CTGTCTGTGGCTGGTGGCTGCCAAGCCTTACCCGAGGGAACGGAGGATAAAGATCAG 600
Db 541 CTGTCTGTGGCTGGTGGCTGCCAAGCCTTACCCGAGGGAACGGAGGATAAAGATCAG 600
QY 601 ACAGCAACGATACCCAGTGTGCTGCGATGCTGCGGCGCCCTCTCTTGTGATGTTCTGG 660
Db 601 ACAGCAACGATACCCAGTGTGCTGCGATGCTGCGGCGCCCTCTCTTGTGATGTTCTGG 660
QY 661 CCAAGTTTCAACTCTGCTGCTGAGAGTCCAAATCGAAGGAAGAAATGCCGTTTCAAC 720
Db 661 CCAAGTTTCAACTCTGCTGCTGAGAGTCCAAATCGAAGGAAGAAATGCCAATGTTCAAC 720
QY 721 ACTTACTATGCTTAGAGTCAAGTGGTGGAGGCAATCTCAGGTCATCTCTTGGCTCAC 780
Db 721 ACTTACTATGCTTAGAGTCAAGTGGTGGAGGCAATCTCAGGTCATCTCTTGGCTCAC 780
QY 781 CCCAAGGGAAGATCAGCAAGCTTATGACAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 840
Db 781 CCCAAGGGAAGATCAGCAAGCTTATGACAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 840
QY 841 GTGGGTACCTCGTGTCACTGATCCCTCTCTCCGTGGCTTGCATGCTGGTGGTGGTGG 900
Db 841 GTGGGTACCTCGTGTCACTGATCCCTCTCTCCGTGGCTTGCATGCTGGTGGTGGTGG 900
QY 901 GTTGGGCTGATCTCCGTGGGGGAGCAAGTACTCTCCGGGGTGTGTAAACCGAGTGGT 960
Db 901 GTTGGGCTGATCTCCGTGGGGGAGCAAGTACTCTCCGGGGTGTGTAAACCGAGTGGT 960
QY 961 GGGATTTCCACAGCTCCATCATGSGGTACACTTCAAGTTCGCTGGTGGTGGTGGAG 1020
Db 961 GGGATTTCCACAGCTCCATCATGSGGTACACTTCAAGTTCGCTGGTGGTGGTGGAG 1020
QY 1021 ATCATCTACATTTGCTGTGCTGTGATACCGTGGAGCCCGCAATGGCATGATGGC 1080
Db 1021 ATCATCTACATTTGCTGTGCTGTGATACCGTGGAGCCCGCAATGGCATGATGGC 1080
QY 1081 TTCCAGTCTCTCTAGCATGGGGAACCTCAGTTGGCAATGTAAGTGTAGCTCTCAGTCT 1140
Db 1081 TTCCAGTCTCTCTAGCATGGGGAACCTCAGTTGGCAATGTAAGTGTAGCTCTCAGTCT 1140
QY 1141 GGTCTCTGACAGTGTGCTCTCTAAATCTTAAATATGGAAGCACTCATGAGGCTAA 1200
Db 1141 GGTCTCTGACAGTGTGCTCTCTAAATCTTAAATATGGAAGCACTCATGAGGCTAA 1200
QY 1201 TATTTTGTGACCAAGTTTCTGGAGTTCCTTCATTTGGCTGTGGATTTTAA 1254
Db 1201 TATTTTGTGACCAAGTTTCTGGAGTTCCTTCATTTGGCTGTGGATTTTAA 1254
```

RESULT 9

AF510069

LOCUS

DEFINITION Homo sapiens Rhesus blood group D antigen D(667) variant mRNA, complete cds.

ACCESSION AF510069

VERSION AF510069.1 GI:25573107

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1480)

Noizat-Pirenne, F., Lee, K., Le Pennec, P. Y., Simon, P., Kazup, P.,

Bachir, D., Rouzaud, A. M., Roussel, M., Juszczak, G., Menanteau, C.,

Rouger, P., Kob, R., Carttron, J. P. and Ansart-Pirenne, H.

Rare RHCE phenotypes in black individuals of Afro-Caribbean origin:

identification and transfusion safety

Blood 100 (12), 4223-4231 (2002)

12393640

2 (bases 1 to 1480)

Noizat-Pirenne, F., Carttron, J. P. and Ansart-Pirenne, H.

Direct Submission

Submitted (08-MAY-2002) CNRGS, INTS, 20 rue Bouvier, Paris 75011,

France

Location/Qualifiers

1..1480

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/note="isolated from Black individual of Afro-Caribbean

origin"

44..1297

/notes="RHD"

/codon_start=1

/product="Rhesus blood group D antigen D(667) variant"

/protein_id="AA075125.1"

/db_xref="GI:25573108"

/translation="MSSKYPERSVRCLPLWALTLEAALLLLFFFFHYDASLEDQKGL

VASYQVQDLTMAAIGLGFISFRRHSWSVAENLFMLAGVQVAILDGLDGLSFPF

SKVLTLSIRLATMSALSVLISVDALGVKNLAQVVMVLVEVTALENDRVLSNI

FNTDYMNMHHIYFAAYFGLSVAMCLPELPEGTEDKDTATISLSAMLGALFMI

FPPNSALLRPIERKNVFNVTYAVSVVTAI SGSLAHPQGIKSTYYSAVLA

GGVAVTGSCHLIPSPWLAMVLGLISVGAKYLPGCCNRVLGPHSIMGYNPSL

LGLGELIYIVLLVLDTVGAGNGMIGFQVLLSIGELSLAIVLALTSGLLTGLLLNLI

WKAPHEAKYFDQVFWKPHLAVGP"

ORIGIN

Query Match

Best Local Similarity

Matches 1252; Conservative

0; Mismatches

2; Indels

0; Gaps

0;

QY 1 ATGAGCTCTAAGTACCGCGGCTCTGCCGCGGTGCTGCCCTCTGGGCCCTTAACTG 60

Db 44 ATGAGCTCTAAGTACCGCGGCTCTGCCGCGGTGCTGCCCTCTGGGCCCTTAACTG 103

QY 61 GAAGCAGCTCTCATCT 120

Db 104 GAAGCAGCTCTCATCT 163

QY 121 CAAAAGGGCTCTGGCATCTTCAAGTGTGCCAAGATCTGACCGTGTATGGCGGCATT 180

Db 164 CAAAAGGGCTCTGGCATCTTCAAGTGTGCCAAGATCTGACCGTGTATGGCGGCATT 223

QY 181 GCGTTGGGCTTCTCCTCCTCGAGTTTCCGGAACACAGCTGGAGCAGTGTGGCTTCAAC 240

Db 224 GCGTTGGGCTTCTCCTCCTCGAGTTTCCGGAACACAGCTGGAGCAGTGTGGCTTCAAC 283

QY 241 CTCCTTCATGCTGGCGTGGTGTGGAGTGGGCAATCTCTGGAGGCTTCTGGAGCCAG 300

Db 284 CTCCTTCATGCTGGCGTGGTGTGGAGTGGGCAATCTCTGGAGGCTTCTGGAGCCAG 343

QY 301 TTCCCTTCTGGGAAGTGGTGTATCACATGTTTCACTTATTCGGCTGGCCACCATGAGTGT 360

Db 344 TTCCCTTCTGGGAAGTGGTGTATCACATGTTTCACTTATTCGGCTGGCCACCATGAGTGT 403

QY 361 TTGTCTGGTCTGATCTCAGTGGATGCTGTCTTGGGGAAGGTCAACTTGGCGCAGTGGTG 420

Db 404 TTGTCTGGTCTGATCTCAGTGGATGCTGTCTTGGGGAAGGTCAACTTGGCGCAGTGGTG 463

QY 421 GTGATGGTCTGGTGGAGGTGACAGCTTTAGCAACCTGAGGATGGTGTATCATATATC 480


```
Db 584 CTGCTCTGGCGCTGTGCTGCTGCCAAGCCTCTACCCGAGGGAACGAGGATAAAGATCAG 643
Qy 601 ACAGCAACGATACCCAGTTTGTCTCCATGCTGGCGGCCCTCTTCTTGTGGATGTTCTGG 660
Db 644 ACAGCAACGATACCCAGTTTGTCTCCATGCTGGCGGCCCTCTTCTTGTGGATGTTCTGG 703
Qy 661 CCAAGTTTCAACTGCTCTGTGTGAGAGTCAANTCGAAGGAAGAAATGCCGTGTTCAAC 720
Db 704 CCAAGTTTCAACTTGTCTGTGTGAGAGTCAANTCGAAGGAAGAAATGCCGTGTTCAAC 763
Qy 721 ACCTACTATGCTGTAGCAGTCAAGCGTGTGTGACAGCATCTCAGGGTCACTCTTGGCTCAC 780
Db 764 ACCTACTATGCTGTAGCAGTCAAGCGTGTGTGACAGCATCTCAGGGTCACTCTTGGCTCAC 823
Qy 781 CCCCAGGGAAGATCAGCAAGACTTATGTGCGACAGTGGCGGTGTTGCGAGGAGCGGTGGCT 840
Db 824 CCCCAGGGAAGATCAGCAAGACTTATGTGCGACAGTGGCGGTGTTGCGAGGAGCGGTGGCT 883
Qy 841 GTGGGTACCTGTGTCACTGATCCCTTCTCCGTGGCTTGCCATGGTGGGTCTTGTG 900
Db 884 GTGGGTACCTGTGTCACTGATCCCTTCTCCGTGGCTTGCCATGGTGGGTCTTGTG 943
Qy 901 GCTGGGCTGATCTCCGTCCGGGGAGCAAGTACCTGCGCGGGTGTGTAAACCGAGTGGT 960
Db 944 GCTGGGCTGATCTCCGTCCGGGGAGCAAGTACCTGCGCGGGTGTGTAAACCGAGTGGT 1003
Qy 961 GGGATTTCCCAAGCTTCAATGAGGCTACAACTTCAGCTTGGGTCTGCTGGAGAG 1020
Db 1004 GGGATTTCCCAAGCTTCAATGAGGCTACAACTTCAGCTTGGGTCTGCTGGAGAG 1063
Qy 1021 ATCATCTACATTTGCTGTGTGCTGTGATACCTGCGAGCGGCAATGGCATGATGGC 1080
Db 1064 ATCATCTACATTTGCTGTGTGCTGTGATACCTGCGAGCGGCAATGGCATGATGGC 1123
Qy 1081 TTCCAGGTCCTCTCCTCAGCAATGGGGAACCTCAGTTGGCCATGCTGATAGCTCTCAGTCT 1140
Db 1124 TTCCAGGTCCTCTCCTCAGCAATGGGGAACCTCAGTTGGCCATGCTGATAGCTCTCAGTCT 1183
Qy 1141 GGTCTCTCAGAGTTTGTCTCTAAATCTTAAATGGAAGACCTCATGAGGCTAA 1200
Db 1184 GGTCTCTCAGAGTTTGTCTCTAAATCTTAAATGGAAGACCTCATGAGGCTAA 1243
Qy 1201 TATTTGTATGACCAAGTTTCTGGAGTTTCTCATTTCCTGCTGGCTGTTGATTTAA 1254
Db 1244 TATTTGTATGACCAAGTTTCTGGAGTTTCTCATTTCCTGCTGGCTGTTGATTTAA 1297
```

```
RESULT 11
AY449381
LOCUS AY449381 1251 bp mRNA linear PRI 03-DEC-2003
DEFINITION Homo sapiens rhesus blood group D antigen DFE mRNA, partial cds.
ACCESSION AY449381
VERSION AY449381.1 GI:38565510
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1251)
AUTHORS Noizat-Pirenne,F. and Ansart-Pirenne,H.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-2003) Blood Bank, EFS, 151 Avenue du Marechal de
Lattre de Tassigny, Creteil 94010, France
FEATURES
source
1..1251
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon.9606"
1..>1251
/notes="weak D DFE"
/codon_start=1
/product="rhesus blood group D antigen DFE"
/protein_id="AAR24078.1"
```

```
/db_xref="GI:38565511"
/translation="MSSKYPRSVRCLPLWALTLEAALLIFYPFTHYDASLEDKGL
VAGYGGDLTVMAALIGLFLTSSFRHSSVAFNLFLMFLADLQVWALLDGLFUSQFP
SGKVIITLFSRLATMSALSVLIISDAVLGKVNLAQLVMVLVEVTAIGNRMVLSNI
PNTDYHNMHLIYVFAAYFGLSVAKLPKPEGTEDKQDTATIPSLSAMLGALFLWI
PWPFSALLRSPIERKNVNTVYAVAVSVTAISGSLAHQPKISKITYVHSAVLA
GDVAVGTSCHLIPSPWLAMVLGLVAGLISVGAKYLPCCNRLVIGIPHSSIMGVNPSL
LGLLGEIYYIVLLVDLTGTAGNGMIGFOVLLSIGELSLAIVALTSGLLTGLLLNLI
WKAPHEAKYFDDQVFWFPHLAVGF"
```

ORIGIN

```
Query Match
Best Local Similarity 99.5%; Score 1247.8; DB 9; Length 1251;
Matches 1249; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGAGCTCTAAGTACCGCGGCTGTGCGGGCGTCCCTGCGCCCTCTGGGCCCTTAACACTG 60
Db 1 ATGAGCTCTAAGTACCGCGGCTGTGCGGGCGTCCCTGCGCCCTCTGGGCCCTTAACACTG 60
Qy 61 GAAGCAGCTCTCATTTCTCTCTTATTTTATCCCACTATGACGCTTCCCTTAGAGAT 120
Db 61 GAAGCAGCTCTCATTTCTCTCTTATTTTATCCCACTATGACGCTTCCCTTAGAGAT 120
Qy 121 CAAAAGGGCGCTCGTGCCATCCCTATCAAGTTGGCCCAAGATCTGACCGTGTATGCGGCCCAT 180
Db 121 CAAAAGGGCGCTCGTGCCATCCCTATCAAGTTGGCCCAAGATCTGACCGTGTATGCGGCCCAT 180
Qy 181 GCGTTGGGCTTCTCACTCGAGTTTCGGAGACACAGCTGGAGCAGTGTGCGCTTCAAC 240
Db 181 GCGTTGGGCTTCTCACTCGAGTTTCGGAGACACAGCTGGAGCAGTGTGCGCTTCAAC 240
Qy 241 CTCCTCATGCTGGCGCTGTGTGCGAGTGGGCAATCTCTGCGAGCGCTTCCTGAGCCAG 300
Db 241 CTCCTCATGCTGGCGCTGTGTGCGAGTGGGCAATCTCTGCGAGCGCTTCCTGAGCCAG 300
Qy 301 TTCCCTTCTGGGAAGTGGTCACTCACTGTTCAGTATTCGGCTGCGCCACCATGAGTGT 360
Db 301 TTCCCTTCTGGGAAGTGGTCACTCACTGTTCAGTATTCGGCTGCGCCACCATGAGTGT 360
Qy 361 TTGTCGGTGTGATCTCAGTGGATGCTGTCTGGGGAAGGTCAACTTGGCGCAGTTGGTG 420
Db 361 TTGTCGGTGTGATCTCAGTGGATGCTGTCTGGGGAAGGTCAACTTGGCGCAGTTGGTG 420
Qy 421 GTGATGCTGCTGTGGAGGTGACAGCTTAGCAACCTCAGGATGCTCATCAGTAAATTC 480
Db 421 GTGATGCTGCTGTGGAGGTGACAGCTTAGCAACCTCAGGATGCTCATCAGTAAATTC 480
Qy 481 TTCAACACAGACTACCAATGAACATGATGCACTCTAGCTTTCGACGCTATTTTGGG 540
Db 481 TTCAACACAGACTACCAATGAACATGATGCACTCTAGCTTTCGACGCTATTTTGGG 540
Qy 541 CTGCTGTGGCTGTGCTGCTGCAAGCCTCTACCCGAGGGAACGAGGATAAAGATCAG 600
Db 541 CTGCTGTGGCTGTGCTGCTGCAAGCCTCTACCCGAGGGAACGAGGATAAAGATCAG 600
Qy 601 ACAGCAACGATACCCAGTTTGTCTGCAAGTCCCAATCGAAGGAAGAAATGCCGTGTTCAAC 720
Db 601 ACAGCAACGATACCCAGTTTGTCTGCAAGTCCCAATCGAAGGAAGAAATGCCGTGTTCAAC 720
Qy 661 CCAAGTTTCAACTCTGCTCTGCTGAGAGTCCAAATCGAAGGAAGAAATGCCGTGTTCAAC 720
Db 661 CCAAGTTTCAACTCTGCTCTGCTGAGAGTCCAAATCGAAGGAAGAAATGCCGTGTTCAAC 720
Qy 721 ACCTACTATGCTGTAGCAGTCAAGCGTGTGTGACAGCATCTCAGGGTCACTCTTGGCTCAC 780
Db 721 ACCTACTATGCTGTAGCAGTCAAGCGTGTGTGACAGCATCTCAGGGTCACTCTTGGCTCAC 780
Qy 781 CCCCAGGGAAGATCAGCAAGACTTATGTGCGAGTGGGTGTTGCGAGGAGCGGTGGCT 840
Db 781 CCCCAGGGAAGATCAGCAAGACTTATGTGCGAGTGGGTGTTGCGAGGAGCGGTGGCT 840
Qy 841 GTGGGTACCTGTGTCACTGATCCCTTCTCCGTGGCTTGCCATGGTGGGTCTTGTG 900
Db 841 GTGGGTACCTGTGTCACTGATCCCTTCTCCGTGGCTTGCCATGGTGGGTCTTGTG 900
```

```
Db      841  GTGGGTACCTCGTGTCACTGATCCCTTCTCCGGCTTGCCATGGTCTGGGTCTTGTG 900
Qy      901  GCTGGGCTGATCTCCGTCGGGGAGCCAAAGTACCTGCGGGGTGTTGTAACCGAGTCTGT 960
Db      901  GCTGGGCTGATCTCCGTCGGGGAGCCAAAGTACCTGCGGGGTGTTGTAACCGAGTCTGT 960
Qy      961  GGGATTCCCAACAGCTCCATCATGGGTACAACTTTCAGCTTGTCTGGGTCTGCTTTGGAGAG 1020
Db      961  GGGATTCCCAACAGCTCCATCATGGGTACAACTTTCAGCTTGTCTGGGTCTGCTTTGGAGAG 1020
Qy      1021  ATCATCTACATGTGCTGCTGCTTGATACCTCGGAGCCGCAATGGCATGATGGC 1080
Db      1021  ATCATCTACATGTGCTGCTGCTTGATACCTCGGAGCCGCAATGGCATGATGGC 1080
Qy      1081  TTCCAGGTCTCCACAGTGGGAACTCAGCTTGGCCATCGTGATAGCTCTCACTCT 1140
Db      1081  TTCCAGGTCTCCACAGTGGGAACTCAGCTTGGCCATCGTGATAGCTCTCACTCT 1140
Qy      1141  GGTCTCCTGACAGGTTTCTCTAAATCTTAAATATGAAAGCACCTCATGAGGCTAAA 1200
Db      1141  GGTCTCCTGACAGGTTTCTCTAAATCTTAAATATGAAAGCACCTCATGAGGCTAAA 1200
Qy      1201  TATTTTGTATGACCAAGTTTCTGGAAGTCTTCTCATTTGGCTGTGGATT 1251
Db      1201  TATTTTGTATGACCAAGTTTCTGGAAGTCTTCTCATTTGGCTGTGGATT 1251

RESULT 12
AY449382      1251 bp  mRNA  linear  PRI 03-DEC-2003
LOCUS      Homo sapiens rhesus blood group D antigen DLO mRNA, partial cds.
DEFINITION
ACCESSION  AY449382
VERSION    AY449382.1  GI:38565512
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 1251)
AUTHORS    Noizat-Pirenne,P. and Ansart-Pirenne,H.
TITLE      Direct Submission
JOURNAL    Submitted (15-OCT-2003) Blood Bank, EFS, 151 Avenue du Marechal de
            Lattre de Tassigny, Creteil 94010, France
FEATURES
            Location/Qualifiers
            source          1..1251
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
            CDS             1..>1251
                        /note="weak D DLO"
                        /codon_start=1
                        /product="rhesus blood group D antigen DLO"
                        /protein_id="AA824079.1"
                        /db_xref="GI:38565513"
                        /translation="MSKYPRSVRRCLPLWALFLEALILLFYFFTHYDASLDQKGL
            VASYQVQDLTVMAATLGLGFLTSFRHSWSVAFNLMLALGVOWAILLDPLSFP
            SGKVITLFSIRLATMSALSVLSVDVLGVKNLAOLVVMVLVEVTALGNLRWVIGNI
            FNTDYHNMWHIYVFAAYFGLSVAVCLPKPLPEGTEDKQDTATIPSLSAMLGALFLWI
            FWPSFNLSLRPIERKNVNTYAYAVSVVTAISGSLAHPQGIKSYVHSAVLIA
            GGVACTLCHLIPSPMLMWLGLVAGLISVGGAKYLPCCNRVLGIPHSISMGYNFSL
            LGLLGLIYIVLLVLDTVGNGWIMIGFQVLLSIGELSLAIVIALTSLGLTGLLLNLKI
            WKAPHEAKYFDQDQVFWKPHLVGF"

ORIGIN
Query Match      99.5%; Score 1247.8; DB 9; Length 1251;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1249; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  ATGAGCTCTAAGTACCGCGGTGTGTCGGCGCTGCTGCCCTCTGGGCCCTTAACACTG 60
Db      1  ATGAGCTCTAAGTACCGCGGTGTGTCGGCGCTGCTGCCCTCTGGGCCCTTAACACTG 60
Qy      61  GAAGCAGCTCTCATCTCCCTCTCTATTTTACCACATGACGCTTCCTTAGAGAT 1200
Db      61  GAAGCAGCTCTCATCTCCCTCTCTATTTTACCACATGACGCTTCCTTAGAGAT 1200
Qy      121  CAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAAAGATCTGACCGTGTGCGGCCCAT 180
Db      121  CAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAAAGATCTGACCGTGTGCGGCCCAT 180
Qy      181  GGGTTGGGCTTCTCAGCTCGAGTTTCGGGAGACACAGCTGGAGCACTGTGGCCCTTCAAC 240
Db      181  GGGTTGGGCTTCTCAGCTCGAGTTTCGGGAGACACAGCTGGAGCACTGTGGCCCTTCAAC 240
Qy      241  CTCTTCATGCTGGCGCTTGTGTGCACTGGCAATCTCTGCGGAGCGCTTCTGAGCCAG 300
Db      241  CTCTTCATGCTGGCGCTTGTGTGCACTGGCAATCTCTGCGGAGCGCTTCTGAGCCAG 300
Qy      301  TTCCCTTCTGGGAAGTGGTCAACACTGTTCAGTATTCGGTGGGCCACCATGAGTCT 360
Db      301  TTCCCTTCTGGGAAGTGGTCAACACTGTTCAGTATTCGGTGGGCCACCATGAGTCT 360
Qy      361  TTGTCGGTGTGATCTCAGTGGATGCTCTTGGGGAAGTCACTTGGCGCAGTTCGTG 420
Db      361  TTGTCGGTGTGATCTCAGTGGATGCTCTTGGGGAAGTCACTTGGCGCAGTTCGTG 420
Qy      421  GTGATGGTGTGTGGAGGTGACAGCTTTAGGCAACCTGAGGATGGTCACTAGTAATC 480
Db      421  GTGATGGTGTGTGGAGGTGACAGCTTTAGGCAACCTGAGGATGGTCACTAGTAATC 480
Qy      481  TTCAACACAGACTACCAATGAAATGATGATGATGATGATGATGATGATGATGATGAT 540
Db      481  TTCAACACAGACTACCAATGAAATGATGATGATGATGATGATGATGATGATGATGAT 540
Qy      541  CTGCTGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db      541  CTGCTGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Qy      601  ACAGCAACGATACCCAGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db      601  ACAGCAACGATACCCAGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy      661  CCAAGTTTCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db      661  CCAAGTTTCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Qy      721  ACCTACTATGCTGTAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db      721  ACCTACTATGCTGTAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy      781  CCCCAAGGGAAGATCAGCAAGACTTATGTGCAAGTGTGCTGCTGCTGCTGCTGCTGCT 840
Db      781  CCCCAAGGGAAGATCAGCAAGACTTATGTGCAAGTGTGCTGCTGCTGCTGCTGCTGCT 840
Qy      841  GTGGGTACCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db      841  GTGGGTACCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Qy      901  GCTGGGCTGATCTCCGTCGGGGAGCCAAAGTACCTGCGGGGTGTTGTAACCGAGTCTG 960
Db      901  GCTGGGCTGATCTCCGTCGGGGAGCCAAAGTACCTGCGGGGTGTTGTAACCGAGTCTG 960
Qy      961  GGGATTCCCAACAGCTCCATCATGGGTACAACTTTCAGCTTGTCTGGGTCTGCTTTGGAGAG 1020
Db      961  GGGATTCCCAACAGCTCCATCATGGGTACAACTTTCAGCTTGTCTGGGTCTGCTTTGGAGAG 1020
Qy      1021  ATCATCTACATGTGCTGCTGCTTGATACCTCGGAGCCGCAATGGCATGATGGC 1080
Db      1021  ATCATCTACATGTGCTGCTGCTTGATACCTCGGAGCCGCAATGGCATGATGGC 1080
Qy      1081  TTCCAGGTCTCCACAGTGGGAACTCAGCTTGGCCATCGTGATAGCTCTCACTCT 1140
Db      1081  TTCCAGGTCTCCACAGTGGGAACTCAGCTTGGCCATCGTGATAGCTCTCACTCT 1140
Qy      1141  GGTCTCCTGACAGGTTTCTCTAAATCTTAAATATGAAAGCACCTCATGAGGCTAAA 1200
Db      1141  GGTCTCCTGACAGGTTTCTCTAAATCTTAAATATGAAAGCACCTCATGAGGCTAAA 1200
```

```
Db 1141 GGTCTCCTGACAGGTTTGTCTCTAAATCTTAAATATGAAGACACCTCATGAGGCTAAA 1200
QY 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTTGGCTTGGATTT 1251
Db 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTTGGCTTGGATTT 1251

RESULT 13
AY449384
LOCUS Homo sapiens rhesus blood group D antigen DLA mRNA, partial cds.
DEFINITION Homo sapiens rhesus blood group D antigen DLA mRNA, partial cds.
ACCESSION AY449384
VERSION AY449384.1 GI:38565514
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1251)
Noizat-Pirenne, F. and Ansart-Pirenne, H.
Direct Submission
Submitted (15-OCT-2003) Blood Bank, BFS, 151 Avenue du Marechal de
Lattre de Tassigny, Creteil 94010, France
JOURNAL
FEATURES
Location/Qualifiers
1..1251
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..->1251
/note="weak D DLA"
/codon_start=1
/product="rhesus blood group D antigen DLA"
/protein_id="AAR24080.1"
/db_xref="GI:38565515"
/translation="MSSKYPKRSVRRCLPLWALTLEAALILFYFFHYDASLEDKGL
VASVQVODLTVAIAIGLGLTSPFRHSWSVAENFMAALGVOWAILLDGFLSOPP
SGKVITLFSRLATMSALSVLISVDVLGKVNLAQLVVMVLVEVTALGNRMVLSNI
ENTYHMMWMLHYFAAFGLSVACLKPLPEGEDKQDQATIPSLSAMGLPFIWI
FWPNSALLSPFERRNAVENTYYAVAVSVVTAISGSLAHPOGKISKTYVHSAVLA
GVAVGTSCPLPSPLAMVLGLVAGLSVGGAKYLPCCNRVLGIPHSNGYNFSL
LGLLEIIVILLVLDITVGAGNMGIFQVLLSIGLSLAIVIALTSLTGLLLNLKI
WKAPHEAKYFDDQVFWKPHLAVGF"

ORIGIN
Query Match 99.5%; Score 1247.8; DB 9; Length 1251;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1249; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTCTAAGTACCGCGGCTCTGCGGGGCTGCGTCCCTCTGGGCCCTAACACTG 60
Db 1 ATGAGCTCTAAGTACCGCGGCTCTGCGGGGCTGCGTCCCTCTGGGCCCTAACACTG 60
QY 61 GAAGCAGCTCTCATTTCTCTCTTCTATTTTATCCCACTATGAGCGTTCTTTAGAGAT 120
Db 61 GAAGCAGCTCTCATTTCTCTCTTCTATTTTATCCCACTATGAGCGTTCTTTAGAGAT 120
QY 121 CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAAAGATCTGACCGTGTATGGCGCCATT 180
Db 121 CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAAAGATCTGACCGTGTATGGCGCCATT 180
QY 181 GGCCTTGGGCTTCTCACCTCGAGTTTTCGGGACACACAGCTGGAGCGTGGCCCTTCAAC 240
Db 181 GGCCTTGGGCTTCTCACCTCGAGTTTTCGGGACACACAGCTGGAGCGTGGCCCTTCAAC 240
QY 241 CTCTTCATGCTGGCGCTTGGTGTGAGTGGGCAATCTCTGCGAGCGGCTTCTTGAGCCAG 300
Db 241 CTCTTCATGCTGGCGCTTGGTGTGAGTGGGCAATCTCTGCGAGCGGCTTCTTGAGCCAG 300
QY 301 TTCCCTTCTGGGAAGTGTGTATCATCTGTTTCAAGTATTCGGCTGGCCACCATGATGCT 360
Db 301 TTCCCTTCTGGGAAGTGTGTATCATCTGTTTCAAGTATTCGGCTGGCCACCATGATGCT 360
QY 361 TTGTCGGTGTGATCTCAGTGGATGCTGTCTTGGGGAAGTCAACTTGGCGAGTTGGTG 420
```

```
Db 361 TTGTCGGTGTGATCTCAGTGGATGCTGTCTTGGGGAAGTCAACTTGGCGAGTTGGTG 420
QY 421 GTGATGGTGGTGGAGGTGACAGCTTTAGGCAACCTGAGGATGCTCATCAGTAATATC 480
Db 421 GTGATGGTGGTGGAGGTGACAGCTTTAGGCAACCTGAGGATGCTCATCAGTAATATC 480
QY 481 TTCAACACAGACTACACATGAAACATGATGACATCTACTACGTTTTCGACGCTATTTGGG 540
Db 481 TTCAACACAGACTACACATGAAACATGATGACATCTACTACGTTTTCGACGCTATTTGGG 540
QY 541 CTGCTGTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 541 CTGCTGTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 601 ACAGCAACATACACGATTTTGTCTGCCATGCTGGGGCGCTCTTCTTCTGGATGTTTCTGG 660
Db 601 ACAGCAACATACACGATTTTGTCTGCCATGCTGGGGCGCTCTTCTTCTGGATGTTTCTGG 660
QY 661 CCAAGTTTCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 661 CCAAGTTTCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 ACCTACTATGCTGTAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 721 ACCTACTATGCTGTAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 781 CCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTGGCGTGTGGCAGGAGGCGTGCT 840
Db 781 CCCCAAGGGAAGATCAGCAAGACTTATGTGCACAGTGGCGTGTGGCAGGAGGCGTGCT 840
QY 841 GTGGGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 841 GTGGGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 901 GCTGGGCTGATCTCCGTGCGGGGAGCCAAAGTACCTGCGCGGGTGTGTAAACGAGTCTG 960
Db 901 GCTGGGCTGATCTCCGTGCGGGGAGCCAAAGTACCTGCGCGGGTGTGTAAACGAGTCTG 960
QY 961 GGGATTTCCCACAGCTCCATCATGCGGCTCAACTTTCAGCTTGTGGGTCTGCTTGGAGAG 1020
Db 961 GGGATTTCCCACAGCTCCATCATGCGGCTCAACTTTCAGCTTGTGGGTCTGCTTGGAGAG 1020
QY 1021 ATCATCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1021 ATCATCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1081 TTCAGGTCCTCTCAGCATTTGGGAACTCAGCTTGGCCATCGTATAGCTCTCACGCTCT 1140
Db 1081 TTCAGGTCCTCTCAGCATTTGGGAACTCAGCTTGGCCATCGTATAGCTCTCACGCTCT 1140
QY 1141 GGTCTCCTGACAGGTTTGTCTCTAAATCTTAAATATGAAGACACCTCATGAGGCTAAA 1200
Db 1141 GGTCTCCTGACAGGTTTGTCTCTAAATCTTAAATATGAAGACACCTCATGAGGCTAAA 1200
QY 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTTGGCTTGGATTT 1251
Db 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTTGGCTTGGATTT 1251

RESULT 14
AY449384
LOCUS Homo sapiens rhesus blood group D antigen DRO mRNA, partial cds.
DEFINITION Homo sapiens rhesus blood group D antigen DRO mRNA, partial cds.
ACCESSION AY449384
VERSION AY449384.1 GI:38565516
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1251)
Noizat-Pirenne, F. and Ansart-Pirenne, H.
```


LGLLGLIIVLVLDVAGAGMIGFOVLISIGELSIAIVIALTSSLITGLLLNLKI
WKAPHEAKYFDDQVFNKPHLAVP"

ORIGIN

Query Match 99.5%; Score 1247.8; DB 9; Length 1251;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1249; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGAGCTCTAAGTACCGCGGTCTGTCGGCGTGTGCTGCCCTCTGGGCGCTTAACACTG	60
Db	1	ATGAGCTCTAAGTACCGCGGTCTGTCGGCGTGTGCTGCCCTCTGGGCGCTTAACACTG	60
QY	61	GAAGCAGCTCTCAATCTCCCTCTCTATTTTACCCACTATGACGCTTCCCTTAGAGGAT	120
Db	61	GAAGCAGCTCTCAATCTCCCTCTCTATTTTACCCACTATGACGCTTCCCTTAGAGGAT	120
QY	121	CAAAAGGGCTCGTGGCATCTCTATCAAGTTGGCCAAAGATCTGACCGTGAATGGCGCAAT	180
Db	121	CAAAAGGGCTCGTGGCATCTCTATCAAGTTGGCCAAAGATCTGACCGTGAATGGCGCAAT	180
QY	181	GGCTTCGGCTTCTCCACCTCGAGTTTCCGAGACACAGCTGGAGCAGTGTGGCTTCAAC	240
Db	181	GGCTTCGGCTTCTCCACCTCGAGTTTCCGAGACACAGCTGGAGCAGTGTGGCTTCAAC	240
QY	241	CTCTTCATGCTGGCGCTTGGTGTGAGTGGGCAATCTGCTGACCGCTTCTTGAGCCAG	300
Db	241	CTCTTCATGCTGGCGCTTGGTGTGAGTGGGCAATCTGCTGACCGCTTCTTGAGCCAG	300
QY	301	TTCCCTTCGGGAGGTGTCTATCACTGTTTCAGTATTTCGGCTGGCCACCATGAGTGT	360
Db	301	TTCCCTTCGGGAGGTGTCTATCACTGTTTCAGTATTTCGGCTGGCCACCATGAGTGT	360
QY	361	TTGTGGGTGCTGATCTCAGTGGATGCTCTCTGGGGAAGTCAACTGGCGAGTTGGTG	420
Db	361	TTGTGGGTGCTGATCTCAGTGGATGCTCTCTGGGGAAGTCAACTGGCGAGTTGGTG	420
QY	421	GTGATGGTGTGGTGGAGTGTGACAGCTTTAGGCAACTGAGGATGTCTATCAGTAATATC	480
Db	421	GTGATGGTGTGGTGGAGTGTGACAGCTTTAGGCAACTGAGGATGTCTATCAGTAATATC	480
QY	481	TTCAACACAGATACCAATGAACATGATGCACATCTACGTGTTCCAGCCCTATTTTGGG	540
Db	481	TTCAACACAGATACCAATGAACATGATGCACATCTACGTGTTCCAGCCCTATTTTGGG	540
QY	541	CTGCTGTGGCGCTGTGCTCCAAAGCCTCTACCCGAGGGAACGAGGATAAAGTCAAG	600
Db	541	CTGCTGTGGCGCTGTGCTCCAAAGCCTCTACCCGAGGGAACGAGGATAAAGTCAAG	600
QY	601	ACAGCAACGATACCCAGTTTGTCTGCATGCTGGGCGCCCTCTTCTTGTGGATGTTCTGG	660
Db	601	ACAGCAACGATACCCAGTTTGTCTGCATGCTGGGCGCCCTCTTCTTGTGGATGTTCTGG	660
QY	661	CCAAGTTTCACTGTCTGTGAGAGTCCATCGAAAGGAAGAAATGCCGTGTTCAAC	720
Db	661	CCAAGTTTCACTGTCTGTGAGAGTCCATCGAAAGGAAGAAATGCCGTGTTCAAC	720
QY	721	ACCTACTATGTGTAGCAGTGTGAGGTCAGCCATCTCAGGGTCATCCTTGGCTCAC	780
Db	721	ACCTACTATGTGTAGCAGTGTGAGGTCAGCCATCTCAGGGTCATCCTTGGCTCAC	780
QY	781	CCCCAAGGGAAGATCAGCAAGACTTATGTGCAAGTGTGGCGGTGTTGGCAGGCGGTGGCT	840
Db	781	CCCCAAGGGAAGATCAGCAAGACTTATGTGCAAGTGTGGCGGTGTTGGCAGGCGGTGGCT	840
QY	841	GTGGGTACTCTGTGTCACTGATGCCCTTCTCCGTGGCTTGCATGGTGTGGGTCTTGTG	900
Db	841	GTGGGTACTCTGTGTCACTGATGCCCTTCTCCGTGGCTTGCATGGTGTGGGTCTTGTG	900
QY	901	GCTGGGTGATCTCGTGGGAGCAAGTACCTGCGGGGTGTGTAAACGAGTGTG	960
Db	901	GCTGGGTGATCTCGTGGGAGCAAGTACCTGCGGGGTGTGTAAACGAGTGTG	960
QY	961	GGGATTTCCCAACAGCTTCCATCATGGGCTACAACCTTCAGTTCGTGGGTCTGCTGGAGAG	1020

Search completed: August 25, 2004, 16:28:51
Job time : 4911 secs

Db	961	GGGATTTCCCAACAGCTTCCATCATGGGCTACAACCTTCAGCTTGTGGTCTGCTTGGAGAG	1020
QY	1021	ATCATCTACATTTGTGCTGTGCTTGTGATACCGTGGAGCCGGCAATGGCATGATTGGC	1080
Db	1021	ATCATCTACATTTGTGCTGTGCTTGTGATACCGTGGAGCCGGCAATGGCATGATTGGC	1080
QY	1081	TTCCAGGTCTCTCTCAGCATTTGGGGAACCTCAGCTTGGCCATCGTGATAGCTCTCACGTCT	1140
Db	1081	TTCCAGGTCTCTCTCAGCATTTGGGGAACCTCAGCTTGGCCATCGTGATAGCTCTCACGTCT	1140
QY	1141	GGTCTCTCAGCAGTTTGTCTCTTAATCTTAAATATGAAAGCACTCATGAGGCTAAA	1200
Db	1141	GGTCTCTCAGCAGTTTGTCTCTTAATCTTAAATATGAAAGCACTCATGAGGCTAAA	1200
QY	1201	TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTTGGATTT	1251
Db	1201	TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTTGGATTT	1251

This Page Blank (uspto)